The Dissertation Committee for Bradley Hall Certifies that this is the approved version of the following dissertation:

COMPUTATIONAL PREDICTION OF ALLOSTERIC NUCLEIC ACIDS

Committee:

Andrew D. Ellington, Supervisor

Robin Gutell

Dmitrii E Makarov

Edward Marcotte

Lauren Meyers

Computational Prediction of Allosteric Nucleic Acids

by

Bradley Hall, B.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2008

Dedication

To my best friend and partner through life, my wife Alisha who through all her support and encouragement helped me achieve my goals. Also, to my "sugar bean," Madison Rose who continuously asked if "I had to work late" and would push me out the door wishing me well at work. Lastly, the push to complete this work was spurred by the imminent arrival of my second daughter, Elizabeth Grace and so to my three girls, I dedicated this dissertation.

Acknowledgements

I should begin by thanking my scientific mentor and advisor Andy Ellington who showed infinite trust and patience with me over the last 9 years. I didn't realize the journey my life would take by moving to Texas and entering his lab, but I would not have changed it for the world. He allowed me the opportunity to travel many scientific paths while always steering me towards success. He also allowed me the opportunity to become involved with the Freshman Research Initiative pilot program, of which begins the next chapter in my life.

I would also like to thank the past and present members of the Ellington lab, including my deskmate, Eun Jeong Cho and the other biosensor designers Manjula Rajendran, Litao Yang and Scott Knudsen for excitement about and knowledge of signaling aptamers.

The lab would not be the same without the utterance of "buddy" at least once a minute. Therefore, I would like to thank the many lab "buddies" including the original buddy, Jay Hesselberth for his involvement in the slip structure work with aptazymes and my love of programming, RNA switches and Linux/Mac. I have fond memories of an afternoon spent shopping for the random coffee maker or beating the crap out of a tennis ball when we were frustrated with code. I would also like to extend a special thanks to the other co-original buddy Matt Levy for most of my bench top skills, including

synthesis and HPLC purification, which were necessary when a dual labeled beacon was required over the weekend for midnight science.

Ram Swamy was instrumental in pushing me to graduate and being a sounding board for all complaints lobbed during my tenure in the Ellington lab, and for this I thank him. In the end, he was a great influence at 3 AM to continue writing another word or crafting another figure. Likewise, research in the Ellington lab would not have been nearly as fun without my buddy Tim Riedel who's enthusiasm for complete off the wall science took him to NASA, Thailand, and currently the University of Southern California. I will cherish our lunch trains on "Easy Fridays", or trips to Posse East, and Kismet. Of course I will never live down the days we would ditch the MURI/Beckman lab to attempt wind surfing on Lake Travis.

I owe a great debt of gratitude to Travis Bayer for his insatiable thirst for science and antiswitch conversations. His dress code brought new meaning to "Causual Fridays". Matt Cowperthwaite was also instrumental for discussions of RNA conformational changes and structure landscapes, while Sean Cater was instrumental in much of the kinetic improvements to the thrombin beacons.

I would also like to thank the many reviewers and those who took the time to read through and edit this and my other professional documents including Eun Jeong Cho, Jessica Silva, Gwen Stovall, Jay Hesselberth, Letha Sooter and Shawn Piasecki. Without their help, my ideas would not have been expressed as clearly on paper as they were in my mind. I would like to extend a special thanks to Shawn Piasecki, who in the short year I've known her, has helped me immensely with the FRI, my research, synthesis and purification, and keeping me organized. The Onion would not have been nearly as fun without her comments.

Nothing I've accomplished would have been possible without my family. I would like to extend a heartfelt and sincere thanks to my parents for their support and understanding that research takes a while and their son will be a Doctor of Philosophy, not a medical doctor or a computer technician. They have supported me mentally, physically and financially through this process and to them I am eternally grateful. Finally, I would like to thank my girls, Alisha, Madison and Elizabeth. You breath fresh life into my world daily and I do this all for you.

COMPUTATIONAL PREDICTION OF ALLOSTERIC NUCLEIC ACIDS

Publication No._____

Bradley Hall, PhD.

The University of Texas at Austin, 2008

Supervisor: Andrew D. Ellington

Selected nucleic acid binding species (aptamers) have been shown to undergo conformational changes in the presence of ligands, and have been adapted to function as biosensors. We were interested in whether the secondary structures of aptamers could be rationally engineered to undergo ligand dependent conformational changes. To this end, we used rational and computational design methods to generate a number of aptamer biosensors.

First, we built upon previous work that showed that antisense oligonucleotides bearing reporter moieties could be used to denature aptamers. Upon addition of ligands, the conformational equilibrium is shifted towards release of the antisense oligonucleotide and a concomitant increase in fluorescence. We attempted to adapt this format to the potential detection of ricin, but were unsuccessful.

In order to better evaluate rational designs, we attempted to use computational modeling methods. Again, aptamer biosensors have previously been engineered based on

ligand-induced reorganization of secondary structure (as opposed to oligonucleotide displacement), a so-called 'slip-structure' model. We developed an algorithm to evaluate different lisp structures, predicted both aptamers and aptazymes that should have undergone ligand-dependent changes in conformation, and experimentally evaluated the computationally predicted sequences. A number of robust biosensors that could respond to the cytokine VegF and the small molecule flavin were discovered.

The computational model was further adapted to an aptamer biosensor that underwent a larger conformational change upon ligand-binding, an antiswitch. In this model, binding of the ligand stabilizes one hairpin structure at the expense of a competing structure (as opposed to merely changing the register of the hairpin as in the previously described slip structure model). Again, we were able to computationally identify a number of antiswitches that upon synthesis were responsive to the ligand theophylline.

Finally we again attempted to use rational design methods to optimize not just the degree of signal but also the kinetic performance of aptamer biosensors. To this end, we developed biosensors that signaled within seconds the presence of the coagulation protein thrombin.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Biosensors are analytical tools capable of transducing analyte detection into easily identifiable optical or electrochemical signals. The ability to sense environmental change is derived from interactions between a target and macromolecule, often leading to changes within the macromolecule. Through engineering, these changes can be converted into a measurable signal representing the presence of a target, even within complex mixtures of nonspecific analytes. Biosensors have become a critical tool in a wide variety of applications including home-based and clinical diagnostics, microbial contamination monitoring, and toxin detection. Most commercially available biosensors today rely on antibody or enzyme recognition. Whereas these methods can be sensitive, there are inherent disadvantages to these schemes including short shelf life, high reagent consumption and complex or expensive detection methods.

Nucleic acids have been discovered that recognize a wide variety of ligands including nucleic acids and more complex non-nucleic acid analytes both in nature (riboswitches) and through *in vitro* manipulation (molecular beacons and aptamers). One of the simplest nucleic acid biosensors is the molecular beacon (Tyagi and Kramer 1996). Molecular beacons are fluorescently labeled oligonucleotide probes that form a stem loop structure in which the loop sequence can bind to complementary nucleic acid sequences. In the absence of the complementary sequence the two ends of the beacon are paired and the fluorescent moieties can interact quenching fluorescence. When the beacon binds its cognate sequence, the two ends are separated by at least the distance of the loop sequence such that excitation of the fluorophore leads to a detectable optical signal. One major limitation of molecular beacons is that they are typically only useful for the detection of

nucleic acid targets and therefore fail as discriminatory biosensors for ions, proteins, and whole cells unless these targets have specific nucleic acid binding properties.

IN VITRO SELECTION OF APTAMERS

Utilizing i*n vitro* selection methods (**Figure 1.1**), nucleic acid binding species (aptamers) have been raised to detect a wide array of different chemical, biological, or physical conditions including ion deficiency, cancer associated proteins, toxins, and whole pathogenic bacteria (reviewed in Nimjee, Rusconi et al. 2005; Proske, Blank et al. 2005; Lee, Stovall et al. 2006). Aptamers are ideally suited for biosensor applications because they undergo ligand-dependent conformational changes, which can be engineered to transduce the recognition event into a detectable signal (Hermann and Patel 2000). They often display binding affinities similar to antibodies, are reproducibly synthesized, can be functionally derivatized and are stable for long periods of time.

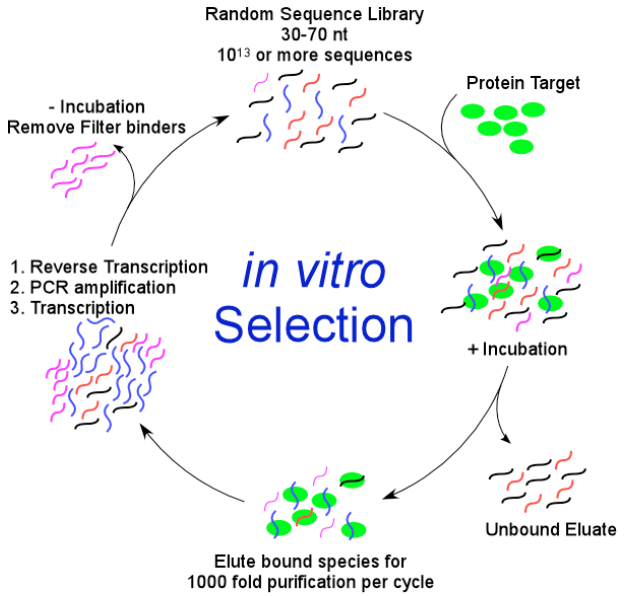


Figure 1.1. Selection schema for the manual selection of aptamer nucleic acids.

The first selected aptamers were single-stranded RNA molecules capable of binding organic dyes (Ellington and Szostak 1990) and T4 DNA polymerase (Tuerk and Gold 1990). To date, many different pools have been utilized for aptamer selection including natural nucleic acids such as ssDNA, dsDNA and RNA or modified nucleic acids including PNA, LNA, 2'OMe RNA and 2'F RNA. In its simplest form, functional nucleic acids are enriched from random sequence populations of chemically synthesized combinatorial oligonucleotide libraries that contain between 20 to 100 random bases. These randomized sequence pools provide ample sequence and structural complexity for high target specificity.

The random sequences are often flanked by constant regions required for the amplification and reconstitution of the pool. These portions can contain primer binding regions for polymerase chain reaction (PCR) or RNA polymerase promoter sequences. Following analyte incubation, functional species are sieved from non-functional species by immobilization on columns or through filtration (**Figure 1.1**). Functional species are then amplified by a combination of reverse transcription, PCR, and *in vitro* transcription depending on the nucleic acid utilized. Multiple cycles of selection and amplification result in the preferential enrichment of those binding species with the highest affinities or activities. Either during the selection or afterwards, the aptamers can be modified for increased stability in potentially hostile environments or can be rationally modified to with enhanced functionality such as biosensing applications.

APTAMER BIOSENSORS

Aptamers can not only be selected against a wide analyte range and therefore act as the biorecognition elements, they can also be engineered to comprise the reporter domain of biosensors, thus introducing the idea of aptamer biosensors (Bier and Furste 1997; Osborne, Matsumura et al. 1997). An aptamer biosensor consists of at least two

domains where the aptamer domain binds to an analyte and one or more signaling sequences that report the presence of an aptamer-analyte complex.

One of the first aptamer-based biosensors was an ELISA-based design where an aptamer replaced the functions of an antibody in analyte detection (reviewed in Jayasena 1999). In this "ELONA", Drolet *et al* developed a vascular endothelial growth factor (VegF) assay by immobilizing the cytokine on microtiter plates and allowing fluorescently labeled anti-VegF aptamers to bind (1996). Anti-fluorescein antibodies were conjugated with alkaline phosphatase and bound to the fluorescein on the aptamers. Through this sandwich type assay, VegF concentration was monitored by the addition of chemiluminescent alkaline phosphatase. These assays still required enzymes and proteins for detection and had similar limitations to conventional ELISA assays.

Since aptamers often undergo conformational changes upon binding their cognate ligands (Burgstaller, Kochoyan et al. 1995; Ye, Gorin et al. 1999; Hermann and Patel 2000; Patel and Suri 2000), it has proven possible to detect binding without the complicated sandwich assays described above. The structural changes have been engineered to report analyte-dependent conformational changes through chemical labeling of various moieties in such a way that function (binding) is not perturbed (**Figure 1.2**). This concept is referred to as a reagentless biosensor where analyte detection is possible without the addition of reagents other than the sample.

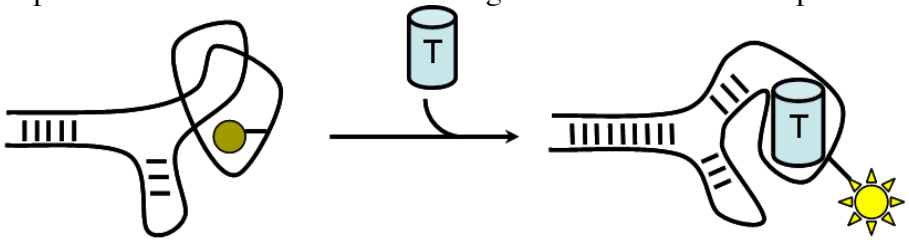


Figure 1.2. Reagentless biosensor that produces a signal upon binding a target

In Vitro Selection of Signaling Aptamers

Designing signaling aptamer biosensors typically follows one of two paths. The first method is to develop the biosensor *de novo* during selection by incorporating both ligand detection and signal generation into the sieving mechanism (Jhaveri, Rajendran et al. 2000; Nutiu and Li 2005; Morse 2007; Rajendran and Ellington 2007). This method is useful when no aptamer has been selected against a particular target of interest or when the a loss of binding affinity is experienced in post-engineered aptamers. It often uniquely combines analyte detection and signaling in a way rational design cannot due to the nature of evolution and selection. However, this process is time consuming and limited to the specific analyte/detection method incorporated and considerable planning must go into the initial design process.

Rational Design of Signaling Aptamers

The second and much more common method is rational design where preexisting aptamers are modified to incorporate some signaling domain. In fact, one of the first signaling aptamer constructs that could directly transduce ligand binding into signal was developed through rational modification of an anti-ATP aptamer (Figure 1.3 A; Jhaveri, Kirby et al. 2000). The authors incorporated a fluorescent moiety into an unstructured internal loop portion of the aptamer, and upon ATP binding the loop was stabilized, altering the environment of a fluorescently labeled adenosine resulting in a concomitant fluorescence change. Various incorporation sites were tested and one was found to provide a significant signal to noise ratio.

Since then, dozens of previously selected aptamers have been converted into signaling aptamer biosensors by the incorporation of various optical (fluorescent or colorimetric) moieties, electrochemical alterations, or by coupling ligand binding to ribozyme activity (summarized in **Table 1.1**).

Signal Strategy	Fluorophore	Fluorescence Change	Signal Mechanism	Target	NA type	Ref
F-labeled sequence	Rhodamine, Cascade blue, Texas red/Base_selection	~80% up	Fluorophore environment change --> quantium yield change --> intensity change	ATP	DNA, RNA	1
	Acridine, Fam/base	~80% up	Fluorophore environment change --> quantium yield change --> intensity change	ATP	DNA, RNA	2
	Bodipy/ribose	1.8-3.5X up	Fluorophore environment change --> quantium yield change --> intensity change	ATP, Tyrosinamide, Argininamide	DNA	3
	Bispyrene/base	3.4X up	Fluorophore environment change --> $F_{eximer\ (480nm)}$ vs $F_{monomer\ (380nm)}$ ratio change	ATP	DNA	4
Ribose Modification	2-amine/FCM	~60% down	Fluorophore environment change --> Reactivity to FCM to 1-amine change --> FCM intensity change	ATP	DNA	5
Purine Modification	2AP, 3MI, 6MAP Fluorescent nucleotide	> 30% up	Fluorophore environment change --> quantium yield change --> intensity change	IgE, thrombin, PDGF	DNA	6
Dye Displacement	$[Ru(phen)_2(dppz)]^{2+}$	~35% down	Intercalation to double helix --> quantium yield change --> intensity change	IgE, thrombin, PDGF	DNA	7
		~55% down		ATP	DNA	8
	Cyanine Dye	~50% down	Dye displacement --> $F_{monomer\ (7600nm)}$ vs $F_{dimer\ (6700nm)}$ ratio change	Cocaine	DNA	9
	Polythiophene/cation polymer	LOD 2 amol	Cationic polythiophene reorientation --> degree of folding change --> intensity change	Thrombin	DNA	10
	MG/aptazyme	8-20X up	Aptamer module catalyze MG aptamer/MG encapsulation --> MG quantium yield change --> intensity change	ATP, FMN, THA	DNA, RNA	11

Table 1.1. Signaling Aptamer Utilizing Induced-Fit Mediated Dye-Reorientation. References included: 1 (Jhaveri, Rajendran et al. 2000), 2 (Jhaveri, Kirby et al. 2000), 3 (Merino and Weeks 2005), 4 (Yamana, Ohtani et al. 2003), 5 (Merino and Weeks 2003), 6 (Katilius, Katiliene et al. 2006), 7 (Jiang, Fang et al. 2004), 8 (Wang, Jiang et al. 2005), 9 (Stojanovic and Landry 2002), 10 (Ho and Leclerc 2004), 11 (Stojanovic and Kolpashchikov 2004). Adapted from (Cho, Lee et al. In Press).

Strategies for Engineering Fluorescent Structure-Switching Aptamer Beacons

Whereas electrochemical detection encompasses a new wave of technologies, fluorescence is currently the most commonly utilized signaling platform (reviewed in Cho, Lee et al. In Press). Fluorescent properties (intensity, wavelength, anisotropy or lifetime) change in response to physical or chemical environment perturbations. Signaling aptamers are referred to as aptamer beacons if they incorporate a pair of fluorescent moieties where signaling is measured by a variation in quenching or fluorescence resonance energy transfer (FRET). The first aptamer beacon was developed

by Hamaguchi *et al* (2001) by modifying the 15-mer thrombin aptamer (Bock, Griffin et al. 1992) with a short oliognucleotide sequence added to the 5' end complementary to the 3' end of the aptamer **(Figure 1.4A (red))**. The most stable structure in the absence of ligand resembled a hairpin loop akin to a molecular beacon. Upon thrombin addition, the hairpin was destabilized in favor of the native ligand binding G-quadruplex structure. When fluorescein and dabcyl were added to the terminal bases of the aptamer, the beacon shifted from a quenched state in the absence of thrombin to a highly fluorescent state in the presence of thrombin (**Figure 1.4A**). The authors tested different extension lengths and found that if the hairpin was too stable, the free energy of ligand binding was not sufficient to destabilize the beacon conformer. However, if the hairpin was too weak, the G-quadruplex would form and relatively high fluorescence was observed in the absence of thrombin. This rationally designed model can be thought of as a refold signaling aptamer due to the refolding of the thrombin aptamer upon ligand binding. Most of the aptamer beacon designs fall into four basic categories explained in the following sections.

Fold/Unfold

The first design paradigm involves an aptamer that is generally destabilized such that the fluorophore and quencher are separated (**Figure 1.3**). Upon ligand binding, the secondary structure is stabilized and the fluorophore is brought within proximity of the quencher. Whereas published designs have shown decreased fluorescence upon binding due to proximity to quenchers (**Figure 1.3B,C**), FRET pairs can be utilized and the ratio of fluorescence from both dyes can be measured. Groups have published fold/unfold signaling aptamers designs around the anti-cocaine (Stojanovic, de Prada et al. 2001), anti-thrombin (Li, Fang et al. 2002) and anti-PDGF (Fang, Sen et al. 2003 respectively).

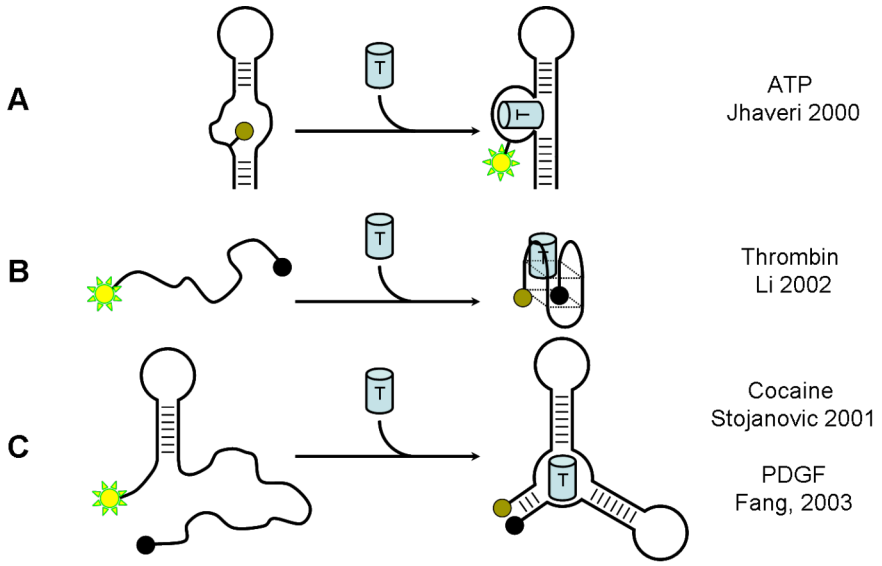


Figure 1.3. "Fold/Unfold" signaling aptamer beacons.

Refold/Displacement

Refold aptamer beacons are more complex in design and involve a global conformational change. Ligand binding involves a restructuring of the secondary or tertiary sequence often separating fluorophore from quencher. As discussed earlier, additional sequence can be added to one or both ends of the aptamer to shift the nonbinding conformer into a molecular beacon structure incorporating both functional moieties into the aptamer itself (**Figure 1.4A**; Hamaguchi, Ellington et al. 2001; Hall, Hesselberth et al. 2007).

Alternatively, additional oligonucleotides can be designed that hybridize to the aptamer and stabilize a nonbinding conformer (**Figure 1.4B**). Nutiu *et al* have designed biosensors incorporating the anti-ATP and anti-Thrombin aptamers in either a bi- or tri-partite model (2003; 2004). Upon ligand binding, the oligonucleotides are denatured yielding an increase in fluorescent intensity.

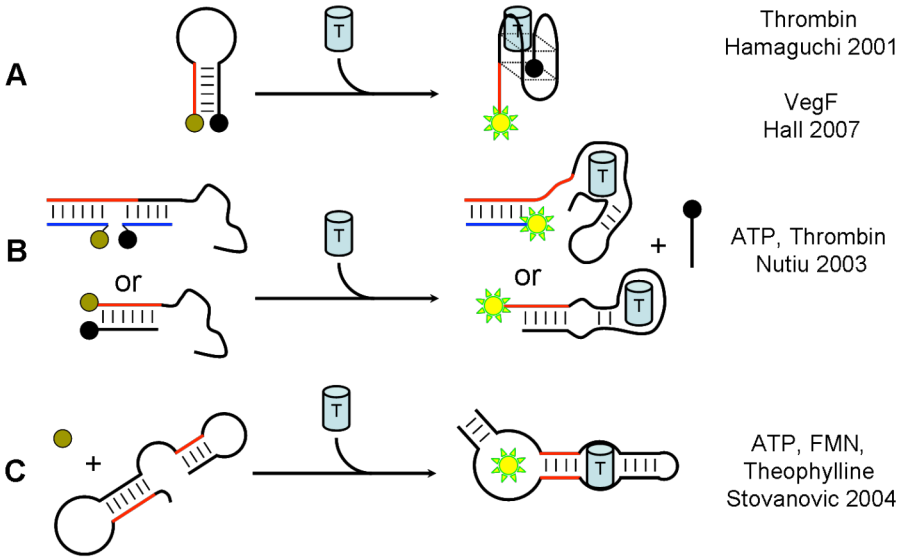


Figure 1.4. "Refold" signaling aptamers.

Lastly, allosteric aptamer sensors have been designed that couple the binding activity of two aptamers where one acts as the "signaling domain" by binding a dye (cibacron blue or malachite green) and the other aptamer is capable of binding the desired target (Wu and Curran 1999; Stojanovic and Kolpashchikov 2004 respectively). In the absence of target, the dye binding aptamer is shifted away from its native structure and incapable of binding the dye. In the presence of target, a structural rearrangement occurs in both aptamers leading to increased fluorescence (**Figure 1.4C**). Similar to previously published allosteric ribozymes (Tang and Breaker 1997; Tang and Breaker 1998; Soukup and Breaker 1999; Robertson and Ellington 2000), Stojanovic *et al* combined an anti-malachite green aptamer with either a cocaine or rATP aptamer (2004). In the absence of ligand, the dye binding aptamer displayed little MG binding and subsequence fluorescence. However, when cocaine or rATP were added, both aptamers bound their corresponding targets and fluorescence increased.

Self Assembly

In certain cases, stable subunits of one aptamer can specifically self associate in the presence of target (**Figure 1.5A**). To design these, an internal loop is often broken to yield two separate sequences. On their own, the sequences do not share sufficient base pairs to form stable hydrogen bonds at the assay temperature, however when the target is added, the free energy from both binding and base pairing stabilizes the aptamer conformer. Both sequence subunits can be labeled with fluorophores (FRET/quenching assay), or one of the subunits can be dual-labeled similar to molecular beacons. This type of biosensor has been developed with the anti-Tat aptamer (Yamamoto, Baba et al. 2000) in addition to the anti-cocaine and anti-ATP aptamers (Stojanovic, de Prada et al. 2000).

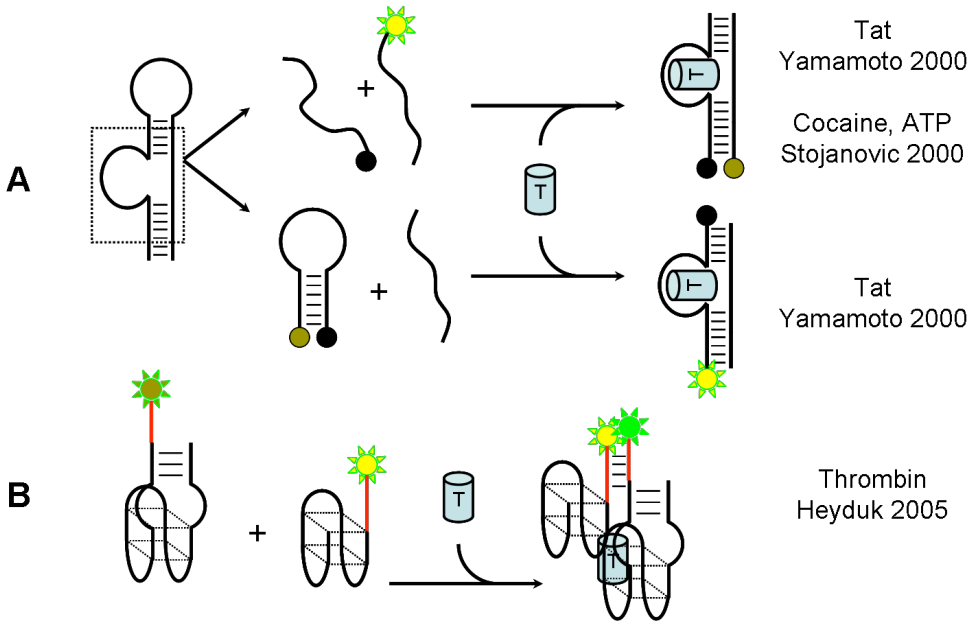


Figure 1.5. "Self Assembly" Signaling aptamer beacon.

Alternatively, multiple aptamers capable of binding different epitopes on the same target can be utilized to bring fluorophores within close proximity in an assay design similar to proximity ligation reactions previously reported (**Figure 1.5B**; Fredriksson, Gullberg et al. 2002; Pai, Ellington et al. 2005). In the absence of target, the aptamers are

not able to associate in solution due to the weak interaction of complementary tails added. However, when they each bind a single target, the proximity greatly reduces the energy needed to self-associate in solution. When fluorophores were added to two different the anti-thrombin aptamers, the addition of thrombin brought the two aptamers close enough that FRET based detection was possible (**Figure 1.5B**; Heyduk and Heyduk 2005).

Assisted Refold/Antiswitch

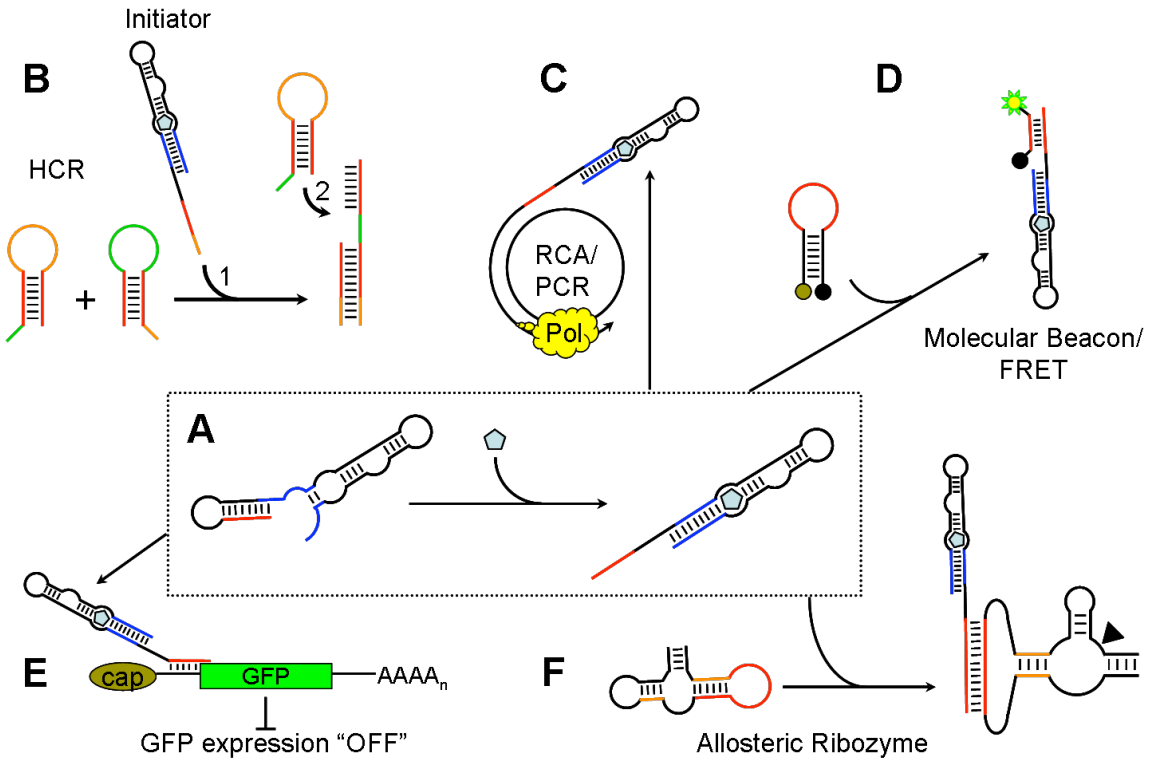


Figure 1.6. An "Antiswitch" aptamer design (A) with various signaling pathways (B-F).

One of the current limitations of signaling aptamer beacons is the lack of a signal amplification cascade upon binding. Since aptamer beacons often display reduced binding to their targets compared with the originally selected aptamer (discussed below), the limit of detection is generally high. Signal amplification schemes could compensate

for this limitation. Typically, one aptamer beacon "signal" is related to one binding event and therefore detection is limited to ligand concentration and the sensitivity of the equipment utilized for detection. A relatively recent and wholly unexplored type of signaling aptamer that may overcome this limitation is the assisted refold or antiswitch developed by Travis Bayer and colleagues with the anti-theophylline or anti-tetracycline aptamers (**Figure 1.6A, E**; Bayer and Smolke 2005). The switch works by forming a stable conformer that traps or "hides" a sequence through hydrogen bonding (**Figure 1.6A** in red) in the absence of ligand. Upon the addition of ligand, an active binding conformer is formed in response to a portion of the aptamer (**Figure 1.6A** in blue) displacing the hidden sequence. This now free sequence is capable of performing a number of different tasks, including activating or inhibiting gene expression *in vivo*, as originally published. It is also capable of acting as a primer for PCR or rolling circle amplification (**Figure 1.6C**; Yang, Fung et al. 2007). Although untested, it could initiate a hybridization chain reaction originally developed by Dirks *et al* (**Figure 1.6 B**; 2004). We provide evidence of antiswitch activation of molecular beacons in this work (**Figure 1.6D**). Additionally, the antiswitch could provide allostery for oligonucleotide dependent ribozymes (**Figure 1.6F**; Penchovsky and Breaker 2005). The possibilities for regulation are vast and can include anything where oligonucleotides are currently utilized today.

Analyte-induced conformational changes.

Each aptamer biosensor listed above underwent analyte-induced conformational changes typically from some unbound "OFF" conformation to a stabilized secondary, tertiary or quaternary structure in the "ON" bound state. There likely exists an equilibrium between binding and nonbinding conformers. The best signaling aptamer should be designed such that in the absence of ligand, the nonbinding conformers are prevalent, whereas, upon ligand addition, the equilibrium shifts to the binding aptamer.

To this end, computational methods can be utilized to predict not only nucleic acid structure but structures poised to switch in the presence of some trigger.

COMPUTATIONAL PREDICTION OF NUCLEIC ACID SECONDARY STRUCTURE

The majority of free energy contained within a nucleic acid molecule resides largely in its secondary (rather than tertiary) structure. This secondary structure is generally composed of Watson-Crick (A-U/T, T/U-A, C-G, G-C) or wobble (G-U, U-G) base pairs. The secondary structure is stabilized in part by hydrogen bonding between the canonical bases and also by "stacking energies" between the aromatic rings (pi-pi interactions; Mignon, Loverix et al. 2005). These stabilizing interactions form between base pairs in helical stem structures irrespective of the actual base. Additionally, non-paired bases are found in loops and bulges. To a lesser extent, tertiary interactions help stabilize the overall structure and can include pseudo-knots, non-standard base pairings and base-triplets (Michel, Ellington et al. 1990; Gautheret, Konings et al. 1994; Leontis, Lescoute et al. 2006).

Nucleic acid secondary structure can be represented in a number of ways, a few of which are drawn in **Figure 1.7**. Each structure contains some subset of common elements called "loops" (recently reviewed in Dirks, Bois et al. 2007). The overall structure can be decomposed into a collection of stacked helical stems (two or more paired bases), interior loops (two paired bases separated by unpaired bases on each side), interior bulges (an interior loop with unpaired bases on only one side), multi-loop branch junction (three or more base pairs separated by unpaired bases) and a hairpin loop (unpaired bases closed by only one base pair) (**Figure 1.7B**).

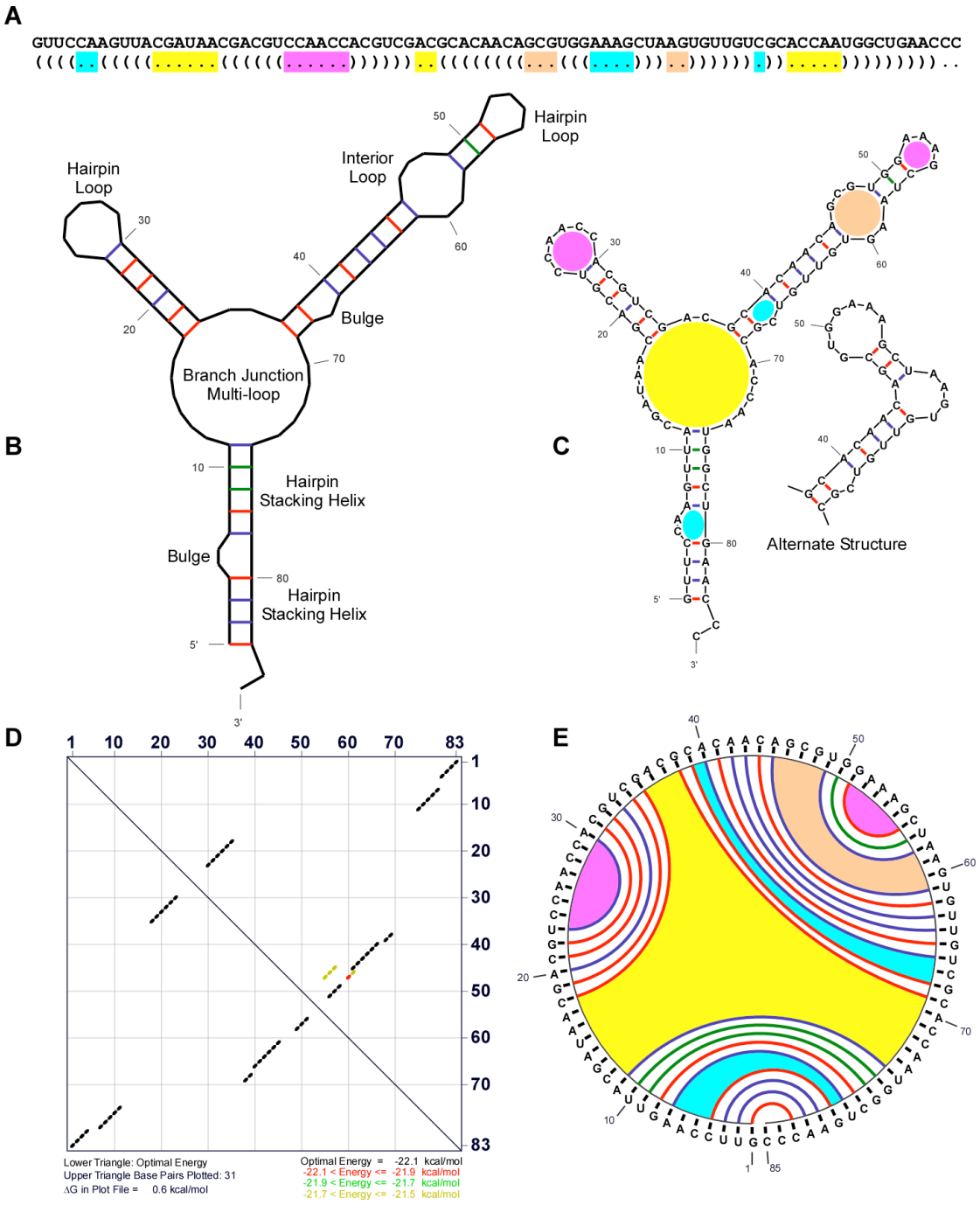


Figure 1.7. RNA Structural Representations from Mfold 3.2. A.) Primary sequence with secondary structure represented in parenthetical notation. Sections are color coded to represent bulges (teal), multi-loop junctions (yellow), hairpin loops (purple) and internal bulges (peach). B.) A generic loop decomposition with common elements of nucleic acid secondary structures. The sequence in (A) is capable of forming this structure. C.) The minimum free energy structure of sequence (A) along with an alternate suboptimal structure of the top right most

stacking helix. Colors and positions are identical to (A). Red, blue and green lines represent potential base pairs. D.) Dot-plot showing base pairs (dots). The lower triangle represents the minimum free energy structure whereas the upper triangle shows potential alternate base pairs. This dot-plot is for the sequence (A). E.) Circle graph representation of the sequence in (A) with identical color coding. It is easy to tell from this graph the structural representation does not contain pseudoknots as the curved lines do not intersect.

Structure prediction by dynamic programming algorithms

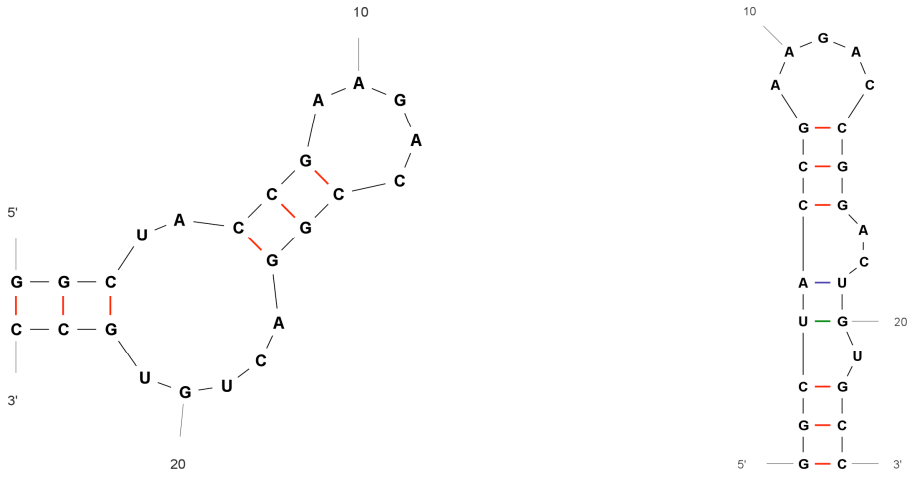
The free energy of a nucleic acid secondary structure is the sum of individual energies of the constituent loops (Tinoco, Uhlenbeck et al. 1971). It is predicted that the number of structures a sequence is capable of forming (N) roughly increases exponentially as a function of length (L) following the equation: $N \approx (1.8)^L$ (Zuker and Sankoff 1984). Even for relatively short sequences (L = 50), there are potentially 5.8E12 different possible structures. If a standard PC could calculate 10,000 structures per second, it would take 280 years to calculate all the possible structures of a sequence 50 bases long. Therefore, the problem is to quickly identify the most common secondary structure represented at equilibrium. For a nucleic acid sequence, this is often the structure with lowest free energy. Using dynamic programming algorithms (DPA) that are capable of implicitly predicting this structure without explicitly traversing over all possible structures is paramount. Ruth Nussinov developed one of the first DPAs to predict nucleic acid secondary structure from its primary sequence nearly 30 years ago (1978). The "Nussinov" algorithm attempted to maximize the number of base pairs using a weighted matrix where G/C was scored 3, A/U(T) scored 2 and G/U(T) scored 1. While fast, this crude method did not take into account many of the loop features that help stabilize or destabilize nucleic acid structures. That same year, Waterman and Smith described a mathematical description of RNA secondary structural components (loops)

and a dynamic programming algorithm to predict these structures (Waterman and Smith 1978).

Soon after, computer programs were developed that incorporated thermodynamic parameters and refined these DPA methods to determine the secondary structure and minimum free energy (MFE) of a sequence (Zuker and Stiegler 1981) (reviewed in: Eddy 2004; Eddy 2004; Mathews and Zuker 2004; Mathews 2006) and the subset of suboptimal structures with the low free energies similar to the MFE (Zuker 1989; Wuchty, Fontana et al. 1999). It is useful to visualize suboptimal structure space as a landscape of peaks and valleys where the most stable structure (with the lowest free energy) is found in the "deepest" valley and there is some energy barrier for the structure to switch conformations into a structure present in nearby valleys. This concept will be expanded on later. Today, nucleic acid secondary structure prediction is generally straightforward using an array of computer programs that utilize the most current thermodynamic parameters and sophisticated DPA approaches (Zuker and Stiegler 1981; Mathews, Sabina et al. 1999; Mathews, Disney et al. 2004).

Decades of research have been funded to experimentally determine thermodynamic "melting" parameters of loop constructs. Most of the work has been published by two main groups: those of Doug Turner (RNA), and John SantaLucia Jr. (DNA) (Walter, Turner et al. 1994; SantaLucia 1998; Xia, SantaLucia et al. 1998; Mathews, Sabina et al. 1999; Mathews and Turner 2002). These empirical nearest-neighbor parameters were discovered by melting specific collections of short double stranded nucleic acids and measuring optical absorbance for the phase transition from double to single strands. Both the change in Gibb's free energy (ΔG) and equilibrium constants (K_{eq}) can be calculated with these parameters. The initial collections represented common structural elements (described in **Figure 1.7B**) with matrices of

closing base pairs that are continually being revised to provide better ΔG estimates. As suggested earlier, the thermodynamic calculation of free energy within a structure is merely a summation of the set of individual loop free energies as described by Tinoco and colleagues (1971). The minimum free energy structure is therefore the structure in the ensemble with the lowest possible set of free energies (**Figure 1.8**).



Structural element	ΔG	Information
External loop	0	0 ss bases & 1 closing helices.
Stack	-3.3	External closing pair is G^{1}-C^{24}
Stack	-3.4	External closing pair is G^{2}-C^{23}
Helix	-6.7	3 base pairs.
Interior loop	3	External closing pair is C^{3}-G^{22}
Stack	-3.3	External closing pair is C^{6}-G^{16}
Stack	-2.4	External closing pair is C^{7}-G^{15}
Helix	-5.7	3 base pairs.
Hairpin loop	4.1	Closing pair is G^{8}-C^{14}
Free Energy	-5.3	kcal/mol

Structural element	ΔG	Information
External loop	0	0 ss bases & 1 closing helices.
Stack	-3.3	External closing pair is G^{1}-C^{24}
Stack	-3.4	External closing pair is G^{2}-C^{23}
Helix	-6.7	3 base pairs.
Bulge loop	1.7	External closing pair is C^{3}-G^{22}
Stack	-1	External closing pair is U^{4}-G^{20}
Helix	-1	2 base pairs.
Bulge loop	3.3	External closing pair is A^{5}-U^{19}
Stack	-3.3	External closing pair is C^{6}-G^{16}
Stack	-2.4	External closing pair is C^{7}-G^{15}
Helix	-5.7	3 base pairs.
Hairpin loop	4.1	Closing pair is G^{8}-C^{14}
Free Energy	-4.3	kcal/mol

Figure 1.8. Calculation of structural free energy. Structural elements and their corresponding free energy as calculated by Mfold server v2.3. The left structure represents the minimum free energy for the sequence, whereas the one on the right is a suboptimal fold with 1 kcal/mol less energy.

It should be noted that the experimental processes utilized to determine the parameters are not complete and free energy structural predictions are therefore limited to

the energy parameters utilized. In addition, tertiary structure and non-Watson-Crick base pairs may affect the predicted secondary structure thus it may not resemble actual structures in biological conditions where the nucleic acid is likely to fold. Whereas free energy predictions are extremely useful, they have limitations due to these and other assumptions and predictive power is often increased by including alternate information such as kinetics of folding, known physical limitations, sequence specific stabilizing energies, etc (Dirks, Lin et al. 2004; Shapiro, Yingling et al. 2007).

Statistical sampling of structural motifs

The MFE structure is often the dominant structural motif occurring at equilibrium for short sequence, however it is often not the only structure a sequence can form at equilibrium. Therefore, statistical methods have been developed to calculate the probability of a given structure within the ensemble of potential structures (McCaskill 1990). Due to the fact that a sequence is able to form many different weak structural interactions, the probability of finding a single structure, even the MFE structure, in the entire ensemble of structures is relatively low. However, most of the lowest free energy structures share structural motifs in common thus enumerating over all possible base pairs is relatively useful (Mathews 2004).

In addition, a DPA has been developed to sample specific suboptimal structures from the complete Boltzmann ensemble of structures (Ding and Lawrence 2003). Bayesian sampling methods have also been developed for entire ensembles of possible structures (Bayesian and statistical methods reviewed in Ding 2006). Using these methods it became possible to calculate the likelihood that certain base pairs will occur and then sample suboptimal structures from an ensemble of structural motifs to find the equilibrium probability of any motif. This overcomes a limitation in the MFE calculation for long sequences, which are structurally difficult to predict due to nearest-neighbor

assumptions built into the model, inaccuracy of energy parameters and tertiary motifs that affect secondary structural arrangements (Doshi, Cannone et al. 2004).

Alternate methods of structure prediction

Structures can be predicted from a given sequences and potential sequences can also be predicted from a given structure (Andronescu, Fejes et al. 2004). If multiple related sequences are available, comparative sequence relationships via covariance models are useful to predict both canonical and non-canonical structural arrangements (Gutell, Lee et al. 2002). Lastly, dynamic programming algorithms have arisen that predict the minimum free energy secondary structure of two or more interacting nucleic acids (Mathews, Burkard et al. 1999; Andronescu, Zhang et al. 2005; Dirks, Bois et al. 2007).

The stochastic folding pathway from linear denatured primary sequence to a stable tertiary or quaternary structure will also influence the overall structural conformation. Denatured RNA follows a kinetic hierarchy when forming its most stable 3-D structure where the secondary structure forms first and then tertiary contacts are formed around this "scaffold" (Hall, Hesselberth et al. 2007) (Brion and Westhof 1997; Tinoco and Bustamante 1999). Recently, software to predict these kinetic pathways has been developed. The Kinfold software suite has been developed to aid predictions of kinetic folding rates based on conformational pathways and includes a visualization tool of folding kinetics (Flamm, Fontana et al. 2000; Xayaphoummine, Bucher et al. 2005). Similary, RNAKinetics has been developed to model the folding kinetics of an elongating RNA, for instance during transcription (Danilova, Pervouchine et al. 2006).

Additionally, non-thermodynamic-based methods have been developed to predict secondary structures. Stochastic context free grammars (SCFGs) are a theoretical way of describing grammar and languages. Since RNA represents a combination of loops and

helix structures, SCFG rule sets have been developed to predict potential RNA structural elements in non-coding RNA (ncRNA) (Rivas and Eddy 1999) and pseudoknot formation (Rivas, Klein et al. 2001) without the utilization of empirically derived free energy parameters .

Software and Resources

Algorithm development is very important to structural prediction, but these algorithms themselves aren't widely useful without implementation through readily available software packages. The two most widely utilized software packages available for nucleic acid secondary structure prediction are the ViennaRNA package (Hofacker 2003) and the MFold related suite of programs (Zuker 2003). Both are available through web interfaces and can be run locally from command line tools. They utilize the same dynamic programming algorithms but differ in the nearest-neighbor parameters for loops (multi-branched and exterior). Additionally, both are in active development with a new release of ViennaRNA (version 1.7) released last month (December 2007).

Some of the other popular computational prediction programs are summarized in **Table 1.2** and a recent benchmark was published, which compared some of their relative performances (Dowell and Eddy 2004). New resources are published through Nucleic Acids Research in its annual "Web Server Issue", now in its fifth year. The Bioinformatics Links Directory, http://bioinformatics.ca/links_directory, contains an actively maintained and comprehensive list of servers and programs for the life sciences that currently includes 47 websites for RNA structure prediction, visualization and design.

Name	Site	Programs	Use	Download or Web Based	Platform
RNASoft	www.rnasoft.ca	PairFold	MFE structure of two input DNA/RNA molecules	Download/Web Based	Linux (Binary/C source)
		CombFold	Tests MFE structure from a combinatorial set		
		RNADesigner	Design RNA sequence with a given secondary structure		
Mfold	frontend.bioinfo.rpi.edu/applications/mfold	Quickfold	MFE and structure for RNA/DNA	UNAFold Download	Linux, Windows binary, MacOSX, C source
		Zipfold	MFE only for RNA/DNA		
		Tm Server	Melt temp for RNA/DNA		
		2-state Hyb	Free energy and Tm for 2+ DNA/RNA sequences		
Vienna RNA	www.tbi.univie.ac.at/RNA	RNAfold	RNA MFE, structure, pair probabilities	Download	Linux (C source)
		RNAeval	MFE		
		RNAheat	Melt curve of RNA sequence		
		RNAinverse	Design RNA sequence with a given secondary structure		
		RNAdistance	Compare secondary structures		
		RNApdist	Compare base pair probabilities		
		RNAsubopt	Compare suboptimal folds		
		RNAplot	RNA structure drawings		
		RNAcofold	Hybridization of two sequences		
		RNAduplex	Predict potential hyb. between two sites.		
		RNAalifold	Covariance to predict consensus structure of multiple sequences		
		RNALfold	predict locally stable structre of long sequences		
		RNAplfold	average pair probabilities for long sequences		
		RNApaln	Fast structural alignment using string alignment		
		Perls Scripts	Useful output utilities		
RNAStructure	rna.urmc.rochester.edu/rnastructure.html	OligoWalk	Predicts binding affinity to RNA targets	Download, OligoWalk available through web interface	Windows Binary
		DynAlign	Predict consensus structure of multiple sequences		
		efn2	Predicts stability of structures		
Pfold	www.daimi.au.dk/~compbio/pfold/	pfold	Predicts common structures from a RNA alignment	Web	HTML Interface
RNA Studio	bibiserv.techfak.uni-bielefeld.de/bibi/Tools_RNA_Studio.html	paRNAss	Alternate RNA secondary structures	Download/Web Based	Linux, Windows binary, MacOSX, C source
		RNA Movies	Visualizing secondary structure landscapes		
		RNAforester	Alignment of RNA trees		
		RNAshapes	Secondary structure probabilities and concensus structures		
		many others			
Sfold	sfold.wadsworth.org/	Srna, others	Statistical predictions for rational design of siRNA, antisense oligonucleotides, trans-cleaving ribozymes, and microRNA	Web Based	HTML Interface
NUPACK	nupack.org	Nupack	Nucleic acid design	Download/Web Based	Linux (Binary/C source)

Table 1.2. Popular nucleic acid structure prediction software.

COMPUTATIONAL PREDICTION OF NUCLEIC ACID CONFORMATIONAL CHANGE

Nucleic acids have been discovered that switch between two or more stable conformations. Examples in nature include transcriptional regulation (riboswitch), translational regulation (bacterial terminator/antiterminator), protein synthesis (EF-Tu/EF-G), mRNA splicing (U6 binding), tRNA maturation, and ribozyme activity (reviewed in Nagel and Pleij 2002; Micura and Hobartner 2003). Likewise, RNA conformational switches have been utilized in the rational design of novel biosensors as previously stated. Until recently however, this rational design was experimentally validated by trial and error of a vast number of potentially viable constructs. Most of the fluorescent aptamer based biosensors described previously required considerable time and effort to test failed constructs. However, there is generally no way to know how many unsuccessful attempts were performed before optimal biosensors were developed because this information is rarely published.

Selection experiments have shown that it is useful to screen and select functional members from a pool of potential candidates. It would therefore seem advantageous to utilize available computation tools to search for sets of potential switching nucleic acids from pools created *in silico* and select only those with predefined characteristics representing an approximate empirical physical model then test the subset experimentally for activity. Using nucleic acid prediction software it has become possible to predict or rationally design potential sequences capable of ligand induced structure switching. Within the last few years, computational methods have been developed to quickly sieve potential structure-switching signaling aptamers *in silico* (Penchovsky and Breaker 2005; Hall, Hesselberth et al. 2007). In 1999, Robert Giegerich presented a method to predict if an RNA sequence is capable of structural switching and wrapped this in a software

package for visualization (paRNAss) that followed three rules (adapted from Giegerich, Haase et al. 1999; Voss, Meyer et al. 2004):

1. There must exist two distinct energy minimum "valleys" representing stable yet significantly different structures such that one must be the global energetic minimum and the other must be found in a structure space with local minima near by.
2. The structures must be separated by an energy barrier to ensure that they exist as two separate structures and the switch between them is only triggered by external events (Flamm, Hofacker et al. 2001).
3. There must not exist other local minima for the RNA to become trapped in once the change has been triggered.

Predicting the possibility of structural switching from known natural switches based on sequence was utilized to proof the algorithms presented by paRNAss. This problem was followed by a method to design a sequence capable of forming predefined structural alternatives (Flamm, Hofacker et al. 2001). Flamm and colleagues presented three examples: (A) create a sequence that consists exclusively of two distinct, predefined structures, (B) create a sequence that is capable of switching structures at a certain temperature and (C) create a sequence with two structures contained in local minima separated by a predefined energy barrier. This last example represents one of the most important factors influencing folding kinetics (Flamm, Fontana et al. 2000). In a simplified form, the energy barrier between two local minimum structures can be calculated by determining the distance in base pairs needed to convert one structure into a less stable structure common to both. This distance measure is paramount to the effectiveness of each model, and various methods are under considerable research. It

should be noted that experimental validation of the predicted sequences presented by Flamm and colleagues was never published.

Researchers in the lab of Danny Barash have also developed computational methods for the design of small riboswitches (Avihoo and Barash 2006) and thermoswitches (Avihoo, Gabdank et al. 2007) using computational methods. Instead of dynamic algorithm approaches that utilize thermodynamic parameters, they incorporate tree-graph theory and have developed eigenvalues to predict switch structures. Unfortunately, the data presented did not provide conclusive evidence that the methods were capable of accurate prediction due to unforeseen structural intermediates and the inability to predict effects of activation energy.

Thermodynamics and kinetics of nucleic acid conformational change

To understand the conformation structure switching designs presented above, it is imperative to provide a basal understanding of energetics and their relationship with respect to free energy and equilibrium constants as they relate to signaling aptamers. For any nucleic acid capable of forming secondary structure, there exists an equilibrium between the unstructured random coil (RC) and some structured states (S_1 or S_2):

(1) $$S_2 \xleftrightarrow{K_2} RC \xleftrightarrow{K_1} S_1$$

where the conversion between states is controlled by the equilibrium constant (K_{eq} or K_1):

(2) $$K_1 = \frac{[S_1]}{[RC]}$$

such that $[S_1]$ is the concentration of structured sequence and [RC] is the concentration of the random coil. The stability of any potential structure formed by a sequence can be quantified by the change in Gibbs free energy from the random coil to a structured state following the equation:

(3) $$\Delta G = \Delta G^\circ_{37} + RT \ln K_1$$

Nucleic acid folding happens at micro- and millisecond time frames so during experimental assays, the structural representations of the sequence is at equilibrium (ΔG = 0) and thus standard Gibbs free energy is utilized. By rearranging equation (3) we derive K_1 as a function of free energy for S_1:

(4) $$K_1 = e^{-(\Delta G^\circ_{37}(S_1)/RT)}$$

corresponds to the change in standard free energy (ΔG°_{37}) of S_1 from the random coil, where R is the gas constant (1.987E-3 kcal/mol/T) and T is the absolute temperature in Kelvin. In nucleic acid structure prediction, T is often 310° Kelvin (37°C). ΔG° is in units of kcal/mol, whereas K a ratio and therefore unitless. Additionally, the difference in structure stability between two structures from the same random coil can be calculated with the following equation:

(5) $$K_{2\to 1} = \frac{K_1}{K_2} = \frac{[S_1]}{[S_2]} = e^{-((\Delta G^\circ_{37}(S_2) - \Delta G^\circ_{37}(S_1))/RT)}$$

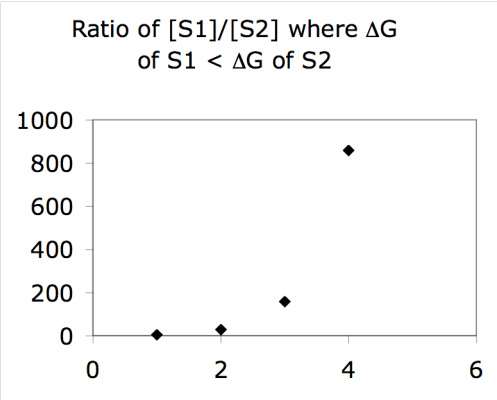


Figure 1.9. Graphical representation of Equation (5). The x-axis is the change in ΔG. The Y-axis corresponds to the molar ratio of more stable to less stable structures.

It is apparent from a graphical representation of equation (5) as shown in **Figure 1.9** that the farther a suboptimal structure's free energy is from the minimum free energy, the less it is represented in solution. A typical cutoff is +3 kcal/mol of the MFE structure

which corresponds to a ratio of 1:150 between structures of X kcal/mol and 3+X cal/mol at equilibrium (Flamm, Hofacker et al. 2001).

As stated earlier, the free energy of the entire structure $\Delta G^\circ_{37}(S_1)$ is calculated by the summation of individual nearest-neighbor loop energies determined experimentally:

(6) $$\Delta G^\circ_{37}(S_1) = \sum_{\text{loop}\in S_1} \Delta G^\circ_{37}(\text{loop})$$

Two state model of structural switching nucleic acids

Signaling aptamers display coupled equilibria between binding and conformational switching. To begin designing aptamer-based biosensors, we must first start with a switching model such as those presented above. For this example, the "refold beacon" switch structure will be utilized. The goal is to take a known aptamer sequence, and append some additional sequence such that it is shifted out of its native binding conformation, i.e. the "aptamer conformer". Upon addition of ligand however, the aptamer conformer should be stabilized such that the structural equilibrium shifts from the non-native conformer to the aptamer conformer. To do this, a sequence is designed such that it incorporates the aptamer sequence and random portion where the minimum free energy structure has paired nucleotides at the 3' and 5' terminus (**Figure 1.10A**) thus forming the non-native structure herein termed the "MFE structure". This provides a quenched state if fluorescent and quenching moieties are derivatized to the ends (green star and red dot in **Figure 1.10A**).

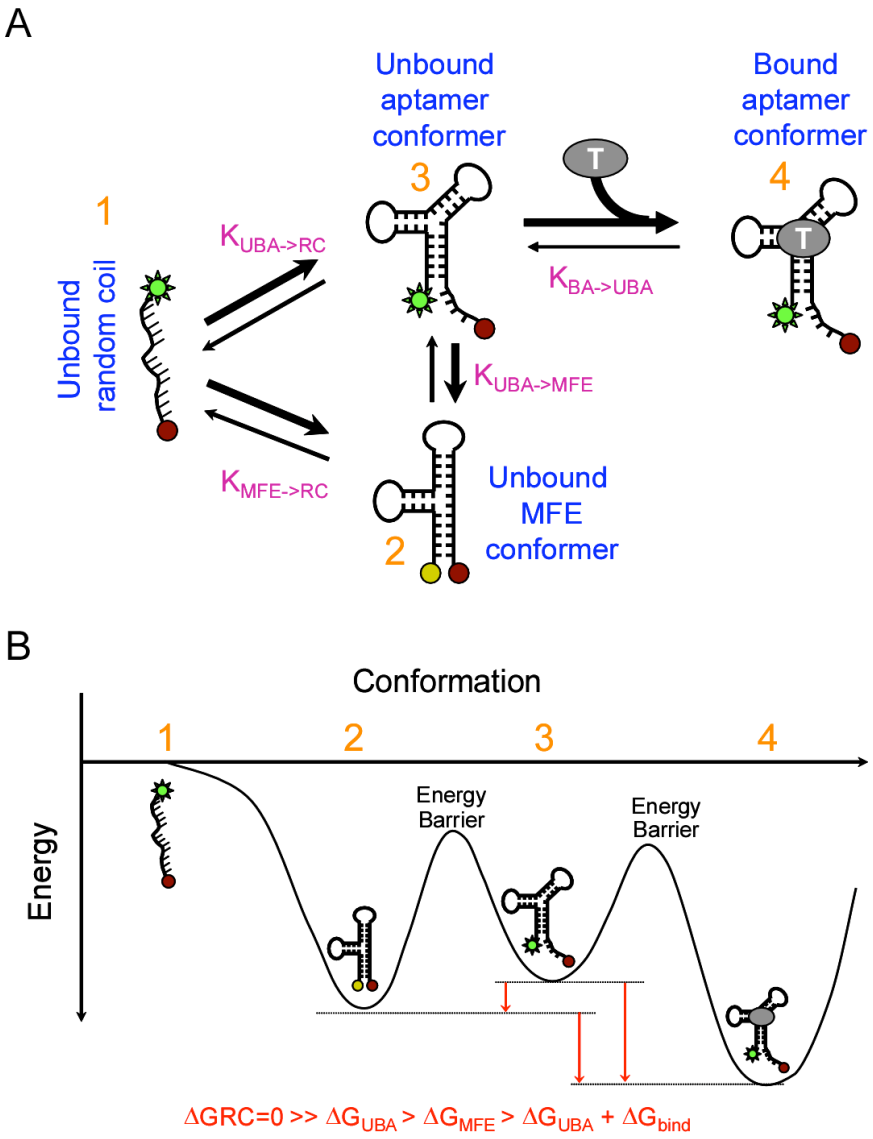


Figure 1.10. Phase diagram for refold aptamer beacon. A.) Possible folding pathways (black arrows) with equilibrium constants (in purple). In a two state model, the aptamer sequence can fold into either an unbound aptamer conformer or unbound MFE conformer depending on the relative thermodynamic stabilities between the two structures. Upon ligand addition, the aptamer conformer should be stabilized shifting the structural equilibrium towards the aptamer conformer. B.) An energy diagram representing the relative free energies of the various conformers (red) and the energy barriers that prevent rapid switching between states.

In a two state model, this sequence should also be capable of forming a second stable structure with a local energy minimum different from the global energy minimum

(as suggested in *Rule 1* by Flamm *et al* (2002) above). This second structure should be capable of binding the analyte, and thus contain the aptamer structure. In the "refold model," the second structure should also shift the terminal nucleotides such that the derivatized fluorophore and quencher are separated, yielding a signal. **Figure 1.10A** represents a phase diagram of this switch whereas **Figure 1.10B** suggests a possible conformation landscape with energy barriers for switching. Both structures may exist at equilibrium in the absence of ligand, however the trigger for switching should be provided by the addition of ligand and is therefore a coupled equilibrium of binding and structure switching.

Equilibrium Analysis of two-state ensemble model

In the two state model, we propose that the aptamer beacon exists in two states in the absence of target in solution: MFE structure and an aptamer conformer. It is unlikely that a significant proportion of molecules exist in the random coil state, but this state is necessary to calculate equilibrium constants. Also, we realize that an ensemble of structures may exist; however we simplify the model for explanation. In the next section we discuss parameters that could be incorporated to account for events such as a structural ensemble.

A concise equilibrium analysis has been proposed for molecular beacons, and we expand it here to incorporate aptamer switching structures (Bonnet, Tyagi et al. 1999). The states and transitions are presented in **Figure 1.10A (blue)** and include a random coil (RC), MFE structure (MFE), unbound aptamer conformer (UBA) and ligand-bound aptamer conformer (BA) with their associated equilibrium constants **(purple)**. The target ligand is represented as T. The fluorescence of the solution at a given temperature is can be explained by the equation:

(7) $$F = \alpha \frac{[\mathrm{BA}]}{\mathrm{A_o}} + \beta \frac{[\mathrm{UBA}]}{\mathrm{A_o}} + \gamma \frac{[\mathrm{MFE}]}{\mathrm{A_o}} + \phi \frac{[\mathrm{RC}]}{\mathrm{A_o}}$$

where α, β, γ, and ϕ, represent the fluorescence of the aptamer beacon in each state and $\mathrm{A_o}$ is the initial concentration of aptamer beacon:

(8) $$\mathrm{A_o} = [\mathrm{RC}] + [\mathrm{MFE}] + [\mathrm{UBA}] + [\mathrm{BA}]$$

The equilibrium constant for the dissociation of structures to random coil can be found by equation (5) as follows:

(9) $$K_{UBA \to MFE} = \frac{K_{MFE \to RC}}{K_{UBA \to RC}} = \frac{[MFE]}{[UBA]} = e^{-((\Delta G^\circ{}_{37}(\mathrm{MFE}) - \Delta G^\circ{}_{37}(\mathrm{UBA}))/RT)}$$

whereas the equilibrium constant for the dissociation of the aptamer-ligand complex can be described by the following equation:

(10) $$K_{\mathrm{BA} \to \mathrm{UBA}} = \frac{[\mathrm{UBA}][\mathrm{T}]}{[\mathrm{BA}]}$$

When ligand is saturating, the total concentration of target $\mathrm{T_o}$ is much greater than the concentration of aptamer beacon $\mathrm{A_o}$, therefore [T] = $\mathrm{T_o}$. Provided the structured aptamer has a sufficiently low ΔG, the equilibrium constant $K_{\mathrm{MFE} \to \mathrm{RC}}$ approaches 0 according to equation (5). This means that in solution, the unstructured aptamer (random coil) is virtually not present. Dropping this term, the fraction of aptamers in each state can be expressed in terms of the target concentration and the two equilibrium constants described above through the equations:

(11) $$\frac{[\mathrm{BA}]}{\mathrm{A_o}} = \frac{\mathrm{T_o}}{\mathrm{T_o} + K_{\mathrm{BA} \to \mathrm{UBA}} + K_{\mathrm{BA} \to \mathrm{UBA}} K_{\mathrm{UBA} \to \mathrm{MFE}}}$$

(12) $$\frac{[\mathrm{UBA}]}{\mathrm{A_o}} = \frac{K_{\mathrm{BA} \to \mathrm{UBA}}}{\mathrm{T_o} + K_{\mathrm{BA} \to \mathrm{UBA}} + K_{\mathrm{BA} \to \mathrm{UBA}} K_{\mathrm{UBA} \to \mathrm{MFE}}}$$

(13) $$\frac{[\mathrm{MFE}]}{\mathrm{A_o}} = \frac{K_{\mathrm{BA} \to \mathrm{UBA}} K_{\mathrm{UBA} \to \mathrm{MFE}}}{\mathrm{T_o} + K_{\mathrm{BA} \to \mathrm{UBA}} + K_{\mathrm{BA} \to \mathrm{UBA}} K_{\mathrm{UBA} \to \mathrm{MFE}}}$$

and substituting these four equations into equation (7) the fluorescence of a solution of

aptamer beacons can be fully described as a function of temperature and ligand concentration by the following equation:

$$F = \frac{\alpha T_o + \beta K_{BA \rightarrow UBA} + \gamma K_{BA \rightarrow UBA} K_{UBA \rightarrow MFE}}{T_o + K_{BA \rightarrow UBA} + K_{BA \rightarrow UBA} K_{UBA \rightarrow MFE}} \quad (14)$$

Thermodynamic substitutions for two-state ensemble model

Each of the equilibrium constants can also be elucidated experimentally through fluorescence measurements as described below using a standard concentration of fluorescent aptamer. The transition of aptamer beacons from the unbound aptamer conformer to the minimum free energy conformer in the absence of ligand ($T_o = 0$), can be calculated by rearrangement of equation (14) as follows:

$$K_{UBA \rightarrow MFE} = \left(\frac{F - \gamma}{\beta - F} \right) \quad (15)$$

where F is the fluorescence of the reaction at equilibrium, β is the characteristic fluorescence of a fully denatured aptamer (found by measuring the fluorescence of aptamer at 80°C) and γ is the characteristic fluorescence intensity of the aptamer in the fully quenched state (this value can either be measured or calculated through the quenching efficiency of the fluorophore:quencher pair (Marras, Kramer et al. 2002)). It should be noted that the aptamer in the ligand binding conformer will have a fluorescence intensity (β) nearly identical to the denatured aptamer (ϕ) if the fluorophore and quencher are sufficiently separated. Also, it should be tested whether ligand-binding acts to quench the fluorophore, which will change the fluorescence intensity of the bound (α) verses unbound (β) aptamer conformers. Both of these conditions can be tested experimentally.

The equilibrium constant for aptamer binding ($K_{BA \rightarrow UBA}$) is analogous to the dissociation constant (k_d) of the aptamer. The dissociation constant represents the ligand

concentration where 50% of the aptamer is bound and 50% of the aptamer is unbound. Quite often however, switch-structure signaling aptamers display higher binding constants than the aptamer alone. Since most switch-structure signaling aptamers are designed to occupy two conformations (active "binding" and inactive "MFE"), only one of these (the active conformer) is capable of binding the target analyte. There is no way to predict *a priori* to what extent sequence modifications will affect the aptamer dissociation constant, however we can adjust for this in the design process by adding a factor (χ) to the k_d of the aptamer:

(17) $$K_{\mathrm{BA}\rightarrow\mathrm{UBA}} = \chi(k_{d\,(\mathrm{aptamer})})$$

and by experimental determination via a ligand binding assay.

Through these equations, it should now be possible to model fluorescence intensity change as a function of protein concentration for any signaling aptamer beacons that switch via a refold mechanism as shown in **Figure 1.10**. Likewise, by modifying the phase model above, it should not prove difficult to predict signaling characteristics of the aforementioned signaling aptamer designs.

Ensemble model of nucleic acid structural landscapes

The ensemble model is similar to the two state model except a Boltzmann distribution of structures at equilibrium replaces the equilibrium constant $K_{\mathrm{UBA}\rightarrow\mathrm{MFE}}$. For instance, if the partition function (described above) is utilized to design a signaling aptamer where the probability of forming the active analyte binding conformer accounts for 2 percent of the structural population, then there is a coupled equilibrium between the two percent bound or unbound and the population dynamic capable of switching between active and inactive structures. In practice, this distribution should be limited to sequences that fall into two separate but similar local energy minima as described by the rules presented by Flamm *et al* (2002). Also, these suboptimal structures should be manually

verified for characteristics that will promote ligand-mediated fluorescent signaling. In the aptamer beacon example above for instance, the population of inactive conformers should consist mostly of structures with paired terminal nucleotides.

CONCLUSIONS

Given the recent advances in nucleic acid structure prediction, developing generic methods to couple analyte binding and optical signaling by target-specific nucleic acid aptamers has become a tractable problem. The focus of this thesis is studying the mechanisms by which an aptamer modulates folding. To this end, generalized methods have been developed to model the secondary structural changes that aptamer-based biosensors undergo. This information was useful to generally relate energetics, structural folds and function. Computational techniques were created to scan large sequence pools for sequences predicted to fold into certain conformers deemed active and experimental validation was provided of their switching capabilities. This ultimately led to the prediction of novel aptamer-based biosensors.

Chapter 2: Tripartite refold aptamer beacon to detect threat agents.

INTRODUCTION

Some of the most toxic substances found on earth are produced naturally by bacteria and plant species. The Merck Index reports the class II ribosome-inactivating protein ricin produced in the castor bean of *Ricinus communis* has a median lethal dose (LD_{50}) of around 1ug ricin per kg body mass whereas the causative agent of botulism, *Clostridium botulinum*, has an LD_{50} in mice of 0.3ng botulin per kg body mass (2006). Thirty years ago the Russian KGB was suspected of using ricin to murder the Bulgarian dissident Georgi Markov by an small prick from an umbrella. A number of aptamers and biosensors have been raised against potential or actual biothreat agents (reviewed in Ligler, Phil et al. 2006). For instance, both RNA and DNA selections have been performed against the ribosome-inactivating proteins ricin (Hesselberth, Miller et al. 2000; Tang, Xie et al. 2006; respectively) and abrin from *Abrus precatorius* (Tang, Yu et al. 2007).

DNA and RNA selections have also been performed against disease causing bacterial targets. A DNA selection against the adhesion protein PilS from *Salmonella enterica serovar Typhi*, the causative agent of typhoid fever, was shown to reduced cell invasion (Pan, Zhang et al. 2005). Additionally, an RNA aptamer raised against the PTPase tyrosine phosphatase was found to inhibit the dephosphorylation of phosphopeptides by Yop51 (Bell, Denu et al. 1998). Yop51 is a virulence determinant in *Yersinia pestis,* which is known to cause bubonic plague.

The anti-abrin aptamer was converted into a signaling aptamer upon ligand dependent binding of the fluorescent dye reporter $[Ru(phen)_2(dppz)]^{2+}$. In addition, aptamer based sensors have been developed to detect ricin utilizing microarray technology (Cho, Collett et al. 2006), so called "electronic tongue" (Kirby, Cho et al. 2004) and capillary electrophoresis (Haes, Giordano et al. 2006). Aptamers frequently undergo small, analyte-induced conformational changes (Hermann and Patel, 2000). By appending fluorescent labels to the aptamer, this change in conformation may be exploited to produce a signal in response to the target. Because the general structure of nucleic acids can be predicted from simple base pairing, it is also possible to engineer aptamers to undergo larger conformational changes in the presence of the target.

In order to design aptamer beacons that respond to biothreat agents, we employed a tripartite strategy in which an antisense oligonucleotide leads to the denaturation of an aptamer in the absence of ligand (**Figure 2.1**; Nutiu and Li 2003). However, in the presence of the cognate analyte, the native aptamer structure forms and the antisense quencher strand is displaced. By appending a sequence to the end of the aptamer complementary to a fluorescently labeled oligonucleotide, target binding (and quencher oligonucleotide release) can be monitored by an increase in fluorescence.

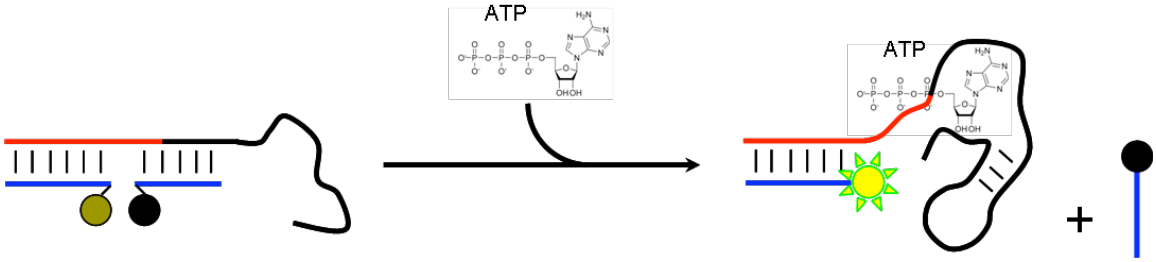


Figure 2.1. Tripartite aptamer beacon construct. The fluorescently labeled oligonucleotide binds to a sequence (red strand) that is appended to the ATP aptamer (black strand). Upon ATP binding, the quencher oligonucleotide is displaced yielding an increase in fluorescence.

The goal was to rationally modify aptamers previously selected to bind ricin (both DNA and RNA) and Yop51 RNA aptamer for signaling in total internal reflection fluorescence (TIRF) microscopy assays. We tested both fluorescent activation via a fluorescent microplate reader and protein binding via filter binding assays. Likewise, various constructs were synthesized to determine the optimal length of antisense quencher and fluorescent oligonucleotides along with distance from the potential core binding sites of the aptamer. Whereas we could reproduce native binding with various constructs and perturb it with various antisense oligonucleotides, none of the rational designs exhibited protein dependent signaling. This was potentially due to poorly elucidated structural requirements within the aptamer necessary for binding, but could also result from the large aptamer size and a lack of ligand induced conformational changes.

RESULTS AND DISCUSSION

Design influences and requirements by TIRF Technologies

This project was initially developed in conjunction with TIRF Technologies (Morrisville, NC) to design aptamer sensors utilized in total internal reflection fluorescent (TIRF) microscopy (**Figure 2.2**). This sensor platform works through evanescent wave penetration across an interface of two materials with different refractive indexes (**Figure 2.2B**). When the excitation wavelength is fully reflected by the first material (typically a glass slide), a highly restricted electromagnetic field identical in wavelength is created within 100nm of the surface of the second material (a flow cell), which can be utilized to excite fluorophores immobilized at the surface (Rumbles, Brown et al. 1991). This method helps to limit the background signal typically present in a fluorescent solution.

To this end, we began designing aptamer beacons that could be immobilized on beads for the detection of biothreat agents through displacement of a quenched moiety leaving behind a fluorescently encoded bead (**Figure 2.2C**). We chose to model aptamer sensors on tripartite designs previously published (Nutiu and Li 2003) and worked with the characteristics of the TIRF platform including attachment chemistries and dye usage. The TIRF Technologies microscope utilizes a 532 nm laser for excitation (**Figure 2.2A**), so initial plans for this project focused on using Alexa532 as the fluorescent reporter and the Iowa Black quencher to best align with the detection capabilities of the TIRF platform. Commercially available organic fluorophores that could be directly incorporated into the aptamer constructs during chemical synthesis were utilized eliminating the need for post-synthesis conjugations and further purification resulting in loss of material.

Another aspect of the design was the difficulty in immobilizing transcribed RNA onto a surface. We hoped to incorporate biotin during *in vitro* transcription of RNA aptamers, as opposed to modifications and purification that may reduce yield after transcription. We agreed on Adegenix's altered biotin-AMP substrate, which could be utilized to prepare 5' biotin-labeled RNA through a single transcription step and high biotinylation yields (70-80%) (Figure 2.3 and Huang, Wang et al. 2003). This allowed aptamer beacon immobilization to streptavadin coated microsphere beads, a condition necessary for the limitations of TIRF microscopy.

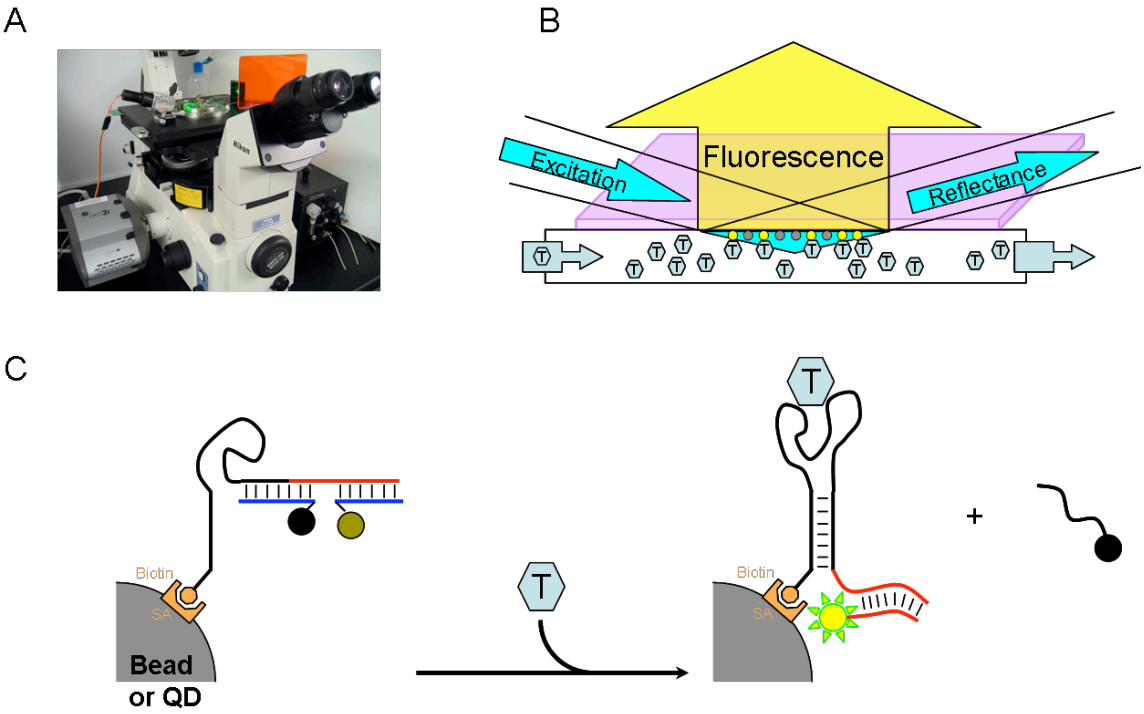


Figure 2.2. Bead immobilized refold aptamer beacon and TIRF microscopy detection. A.) TIRF microscope B.) Diagram of flow cell detection of analyte (T) by bead immobilized aptamer (inactive: gray circle or active: yellow circle). The excitation wavelength is fully refracted, yet an evanescence wave travels ~100nm into the flow cell to excite the fluoresceinated aptamer. C.) Bead level diagram showing activation when target analyte binds and displaces the quencher oligonucleotide.

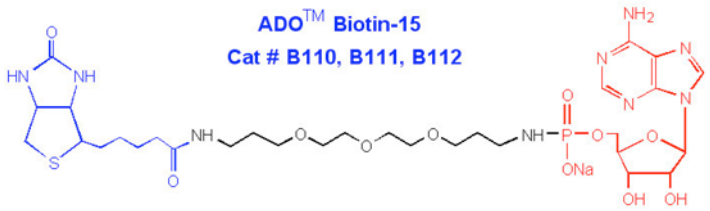
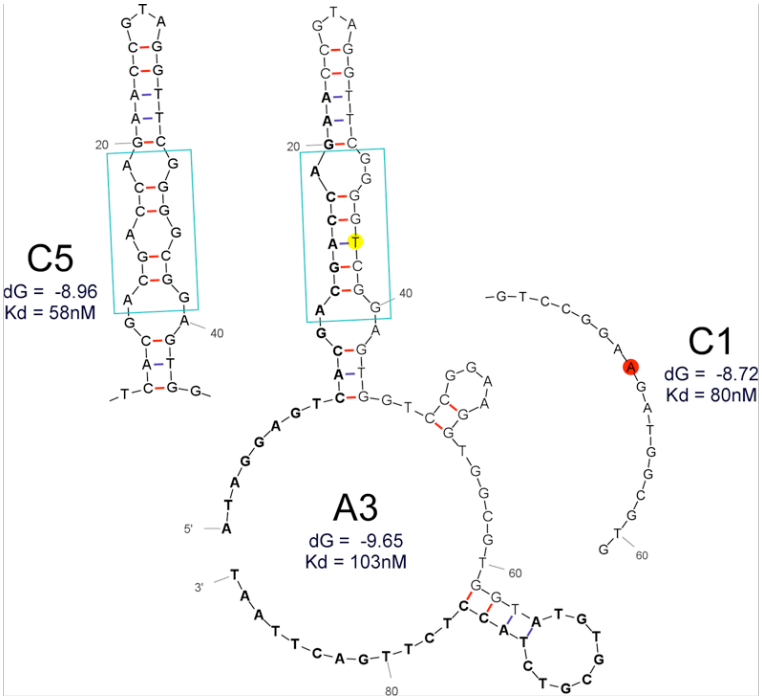


Figure 2.3. Altered biotin incorporation via transcription. Adegenix's altered biotin-AMP incorporated with the following T7 RNA polymerase binding site: TAATACGACTCACTATT**A**GG. The bold A is the first RNA base incorporated during transcription representing the biotin-AMP .

Anti-ricin DNA signaling aptamer design

Initially we set out to design aptamer beacons around DNA aptamers selected against ricin (Tang, Xie et al. 2006). The original aptamers were selected against the entire ricin dimer (A and B chains). However, in order to carry out both engineering and TIRF-based detection under BSL 2 laboratory conditions (or lower) we only utilized the commercially available A-chain for assays. RTA is incapable of penetrating cellular surfaces without the addition of the B-chain (Lord, Hartley et al. 1991).

We began with sequence C5 from a number of possible candidate sequences nanomolar binding affinities(Tang, Xie et al. 2006). Because the sequence had not been well-characterized, we began by making a series of deletion constructs to elucidate the most important structural motifs required for binding. We designed and synthesized 4 different deletion constructs in addition to the original C5 aptamer and tested them for binding (**Figure 2.4B**). C5.2 and C5.3 were designed to test whether ricin bound and stabilized the hairpins into a four-way junction similar to the PDGF and cocaine signaling aptamers (**Figure 2.4A**; Stojanovic, de Prada et al. 2001; Fang, Sen et al. 2003 respectively). C5.3 and C5.4 represented minimum aptamer sequences that would stabilize a potential G-quadruplex to determine if it was responsible for binding similar to the anti-thrombin aptamers (Bock, Griffin et al. 1992).




```
C5      ATAGGAGTCACGACGACCAGAACCGTAGGTTCGGGGCGGAGTGGTCCGGAAGGTGGCGTGGTATGTGCGTCTACCTCTTGACTTAAT
C5.1    TTGGGGTTCACGACGACCAGAACCGTAGGTTCGGGGCGGAGTGGTCCGGAAGGTGGCGTGGTATGTGCGTCTACCTCCCCCC
C5.2    TGGGGGTTCACGACGACCAGAACCGTAGGTTCGGGGCGGAGTGGTCCGGAAGGTGGCGTGGTATGTGCGTCTACCTCCCCCC
C5.3          CACGACGACCAGAACCGTAGGTTCGGGGCGGAGTGG
C5.4         TCACGACGACCAGAACCGTAGGTTCGGGGCGGAGTGGTCCGGAAGGTGGCGTGGTATGT
```

Figure 2.4. DNA binding aptamers for ricin. Bold sequence denotes primer binding sequences for PCR amplification. G to A transition (red circle) in clone C1 destabilizes the small hairpin. Thymine deletion in C5 (yellow circle) changes loop dynamics. C5 has a Kd of 58nM, C1 displays a Kd of 80nM whereas A3 has a Kd of 105nM. The small hairpins near the 3' end were not stable in the subpopulation of sequences. B.) C5 ricin aptamer sequence and 4 modified sequences. Bold sequence represents the primer binding sites as above, blue sequence is a duplex that could potentially stabilize the terminal nucleotides.

Unfortunately, we were unable to detect binding with the original C5 aptamer or any of the deletion/mutation constructs (data not shown) suggesting that the original aptamer was specific for either the entire ricin dimer or the B chain alone.

Anti-ricin RNA signaling aptamer design

It would have been preferable to work with a DNA aptamer due to the ease with which constructs can be synthesized, however the lack of detectable protein binding

prevented its utilization. We therefore began developing sensors based on anti-ricin RNA aptamers originally developed in our lab by Jay Hesselberth (Hesselberth, Miller et al. 2000). One of the published aptamer sequences was found to bind ricin A-chain (RTA) with high affinity (K_d of 7.4 nM +/- 1.1 nM) and selectivity. This aptamer sequence contained a GAGA known to be universially conserved in eukaryotic 23-28 S rRNA and to which RTA was found to be catalytically active (Chen, Link et al. 1998). However, this sequence is capable of forming a number of strucurally diverse species based on computational folding by free energy minimization. Therefore, Hesselberth and colleagues tested two minimal sequence variants that either stabilized a structure in which the GAGA tetramer was sequestered in a stem (RA80.1.d1) or another in which the GAGA tetramer was displayed in a hairpin loop (RA80.2.d1). The authors hypothesized that the exposed GAGA would provide tighter ricin binding since it resembled the naturally occuring tetraloop. Their data however, suggested that RA80.1.d1 was infact the most likely structural variant able to binding RTA even though the GAGA tetramer was sequestered. They then further minimized this sequence (31RA) to detect whether the GAGA was necessesary for recognition and binding or could be removed entirely. This new aptamer bound with a K_d 30 fold lower than the original unmodified aptamer suggesting the tetramer was not imporant. It was noted however that a purine rich GGGG tetramer was displayed in a hairpin loop and this may be sufficient for tight binding.

Since the longer aptamer was capable of forming a number of conformational variants and only a handful of them were shown to bind ricin we incorporated this potential conformational switching into designs. We designed a number of constructs with various stem lenghts and compared binding of these to the original and minimal published aptamers. The full-length aptamer RA80.1 and the minimal aptamer 31RA,

were transcribed and binding was confirmed by nitrocellulose filter binding assays in the presence or absence of RTA. The percent binding was calculated as the fraction of RNA, which could bind the nitrocellulose filter. This is typically very low in the absence of protein target and increases when the protein is present. Both RA80.1 and 31RA could be captured at equimolar concentrations of RNA:protein (**Figure 2.5, left two constructs**) however to a lesser extent than previously published.

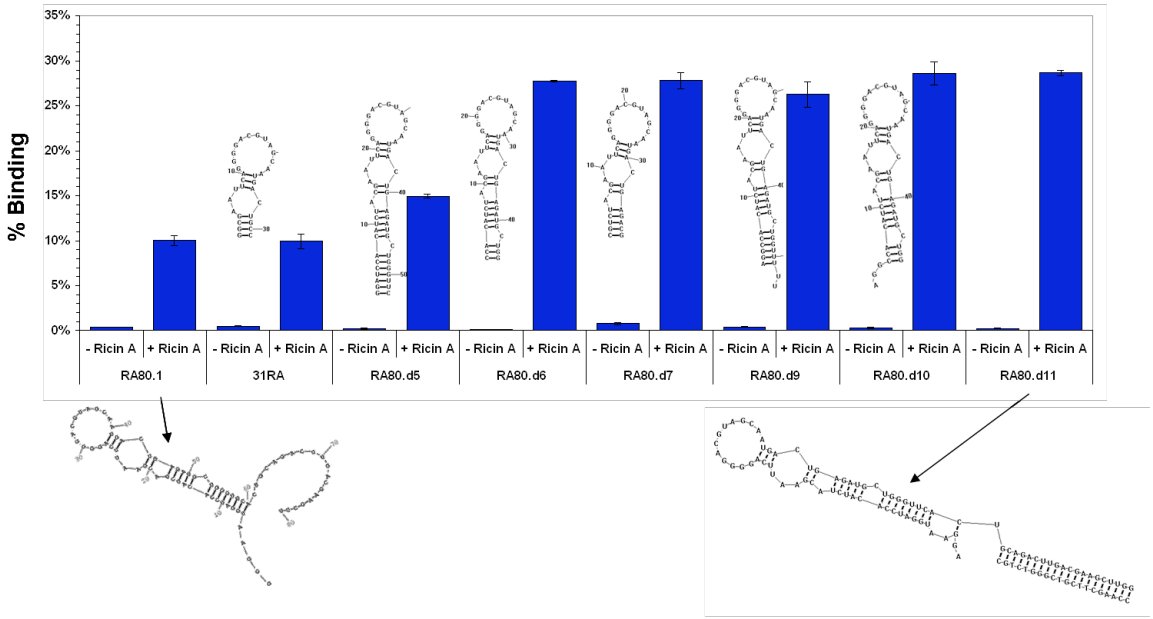


Figure 2.5. RNA ricin aptamer binding assay. The structural conformations of each construct are shown above (or below) the respective binding assay in the presence or absence of RTA.

Based on these results we began to model sequence modifications for the conversion of the aptmers into signaling aptamers using tripartite refold designs described previously (**Figure 2.2 C; Nutiu and Li 2003**). In short, we wanted to develop a signaling aptamer that in the absence of protein effector would stably hybridize to quencher- and fluorophore-labeled oligonucleotides, but in the presence of the protein effector the aptamer would displace the quencher oligonucleotide, ultimately resulting in

an increase in fluorescence. A series of constructs (RA.80.d5 through RA80.d11) were designed and made. RA80.d11 also contained the sequence AGGA at its 5' end required for transcription and incorporation of biotin derivatized adenosine, however as a control, this biotin was not incorporated for the binding or activation assays such that native binding could be determined. Binding assays of all modified sequences were performed and (surprisingly) they yielded enhanced ricin-dependent binding with little to no increase in filter binding alone (**Figure 2.5**). Additionally, we found rescued binding by including an oligonucleotide that would hybridize to the 3' portion of RA80.d11 thus stabilizing the predicted binding structure (compare RA80.1 with RA80.d11, **Figure 2.5**). Other experiments (not shown) confirmed the specificity of binding.

A number of quencher oligonucleotides were designed with predicted melting temperatures and sequence length characteristics similar to those originally designed by Nitiu *et al* such that they would disrupt the aptamer secondary structure (**Figure 2.6A**; red sequences). We hoped they would also display similar activation levels.

As can be seen in (**Figure 2.6B red and green bars**), no ricin dependent activation was detected. The maximum fluorescence in which no quencher was present is about 9-fold greater than in the presence of the longest quencher Q0 and quenching was dependent on the presence of the RNA aptamer (compare **No Q,** and **Q0** with or with out **RA80.d11** in **Figure 2.6B**). Whereas the longest quencher (Q0) presumably bound and quenched fluorescence, when the sequence length was shortened the quenching efficiency also reduced to the point in which the quencher presumably no longer efficiently hybridized to the aptamer.

To ensure that the aptamer was capable of binding ricin in the presence of the added fluorescent oligonucleotides, a binding assay was performed with radiolabeled aptamer constructs. **Figure 2.6C** shows that F0 does not inhibit binding, however Q0

does partially inhibit binding compared to Q3 but does not completely abolish it. This is in line with the predictions based on our knowledge of structure-function relationships in the aptamer. However, the inhibition of binding by Q0 is relatively small, and thus the potential for signal generation is also small.

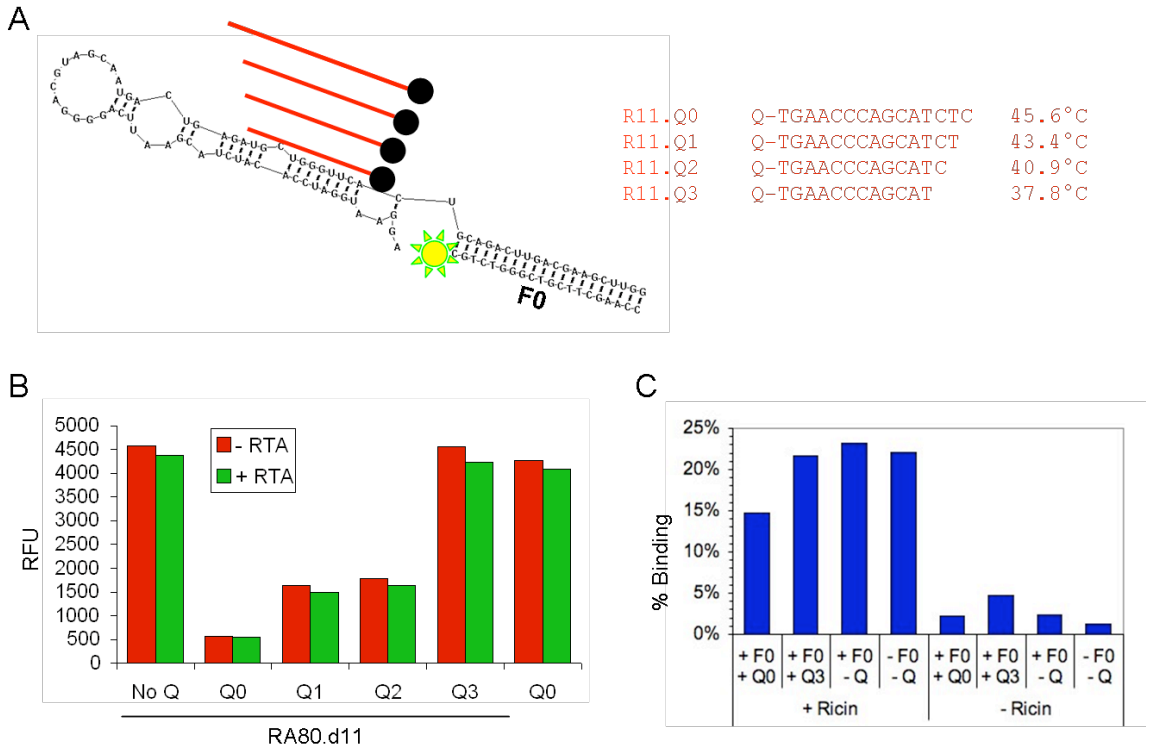


Figure 2.6. Sequence designs and assays for RA80.d11. A.) RA80.d11 structure with F0 oligo hybridized. Red bars indicate relative placement of the red quencher sequences. Also included are the predicted melting temperatures for each oligo. B.) Activation assay of RA80.d11 with F0 and different quencher constructs. C.) Binding assay of two quencher constructs in the presence or abasence of ricin.

We reasoned that Q0 was sufficiently stable under assay conditions and thus difficult to displace. Also, since Q3 did not display quenching, the 5' portion of the aptamer was capable of forming a more stable duplex with itself than the antisense quencher oligonucleotide. Displacement could be achieved by either increasing the assay temperature to universally destabilize hydrogen bonding or shorten the 5' portion of the

aptamer and utilize one of the shorter antisense quenchers. It was necessary to reduce observed competition without increasing the assay temperature since this would also destabilize potentially imporant structures required for aptamer-ligand interactions.

Following these considerations, several deletion variants of the original RA80.d11 were made in which 2, 4, 6, or 8 residues were removed from the 5' portion of the hairpin stem. When 2 bases were removed and the aptamer was tested with F0 and either Q0 or Q3, a reduction in binding was observed. However, when 4 bases were removed, there was no observable difference in binding activity with either Q0 or Q3 (**Figure 2.7**). In the absence of RTA, background binding was consisiently below 3% (data not shown).

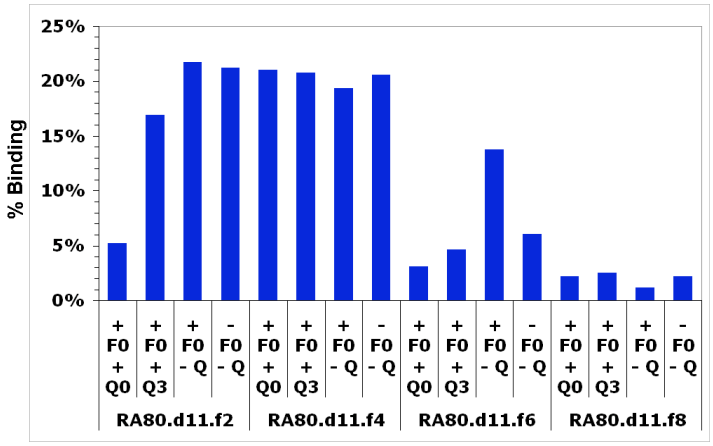


Figure 2.7. Binding assay in the presence of antisense quencher Q0 or Q3 and fluorescent oligonucleotide. RNA was RA80.d11 with 2, 4, 6 and 8 bases removed from the 5' end of the aptamer.

When 6 bases were removed from the 5' end of RA80.d11, ricin-specific binding was partially lost, but could rescued by the addition of the fluorescent oligonucleotide. This would suggest the sequence predominatly forms an inactive conformer that prohibits ricin binding, but when a portion of the sequence is sequestered by the fluorescent oligonucleotide, the native binding conformer prevails (**Figure 2.7; RA80.d11.f6**). When 8 bases were removed, binding was inhibited irrespective of the presence of F0

(**Figure 2.7; RA80.d11.f8**). We propose that destabilization of the stem structure in fact poised the aptamer for a conformational change related to ricin-binding. When the quencher oligonucleotides were added to RA80.d11.f6, they also inhibited ricin binding, consistent with this interpretation. Indeed, now Q3 as well as Q0 could suppress ricin-binding.

The various deletion constructs were also assayed for ricin dependent fluorescence increases, however, no fluorescent changes were observed (**Figure 2.8**). It again appeared that whereas the oligonucleotides are influencing the binding ability of the aptamer, there was no detectable ricin dependent conformational switching between quenched and unquenched states. Unlike previous constructus however, in which the quencher oligonucleotide shows less quenching ability (**Figure 2.8**: RA80.d11-f4 Q3) full ricin binding is observed.

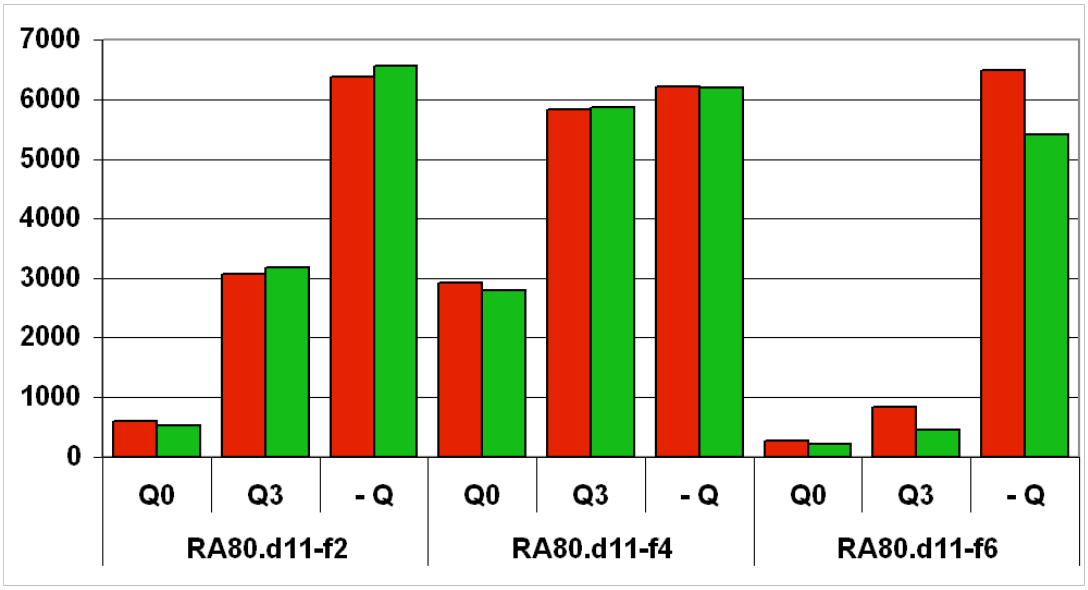


Figure 2.8. Activation assay of 5' aptamer deletion constructs in the absence (red) or presence (green) of 1.33uM RTA after 30 minutes.

Anti-YOP signaling aptamer designs

Ultimately, we did not observe ricin dependent activation although we could inhibit binding with stable quencher oligonucleotides or reduce quenching with unstable quencher oligonucleotides. Therefore, we decided to switch to a different aptamer known to bind biological targets of interest.

To this end, we began working with an anti-tyrosine phosphatase (PTPase) RNA aptamer previously developed in our lab (Bell, Denu et al. 1998). The phosphtatase is a product of the Yop51 gene of *Yersinia enterocolitica*, a close relative of the pathogen *Yersinia pestis*. We incorporated the anti-Yop51 aptamer N71yc2 into our experiments due low nanomolar binding affinities similar to the anti-ricin RNA aptamer (18 nM).

Using the same methods as previously described, we designed a number of antisense quenchers (Yc.2.Q1-3), and an extended the aptamer (N71yc2.d2) with the additional 3' sequence capable of hybridizing to a fluorescently labeled oligonucleotide (Yc.2.F1) (**Figure 2.9A**). We again performed a binding assay to show that the presence of the additional fluorescent oligonucleotide did not prevent the aptamer from binding Yop51 (**Figure 2.9B**). As was the case with the anti-ricin aptamer the quenching oligonucleotides do not appear to inhibit Yop51 (**Figure 2.9B**) and in fact their presence restores full binding.

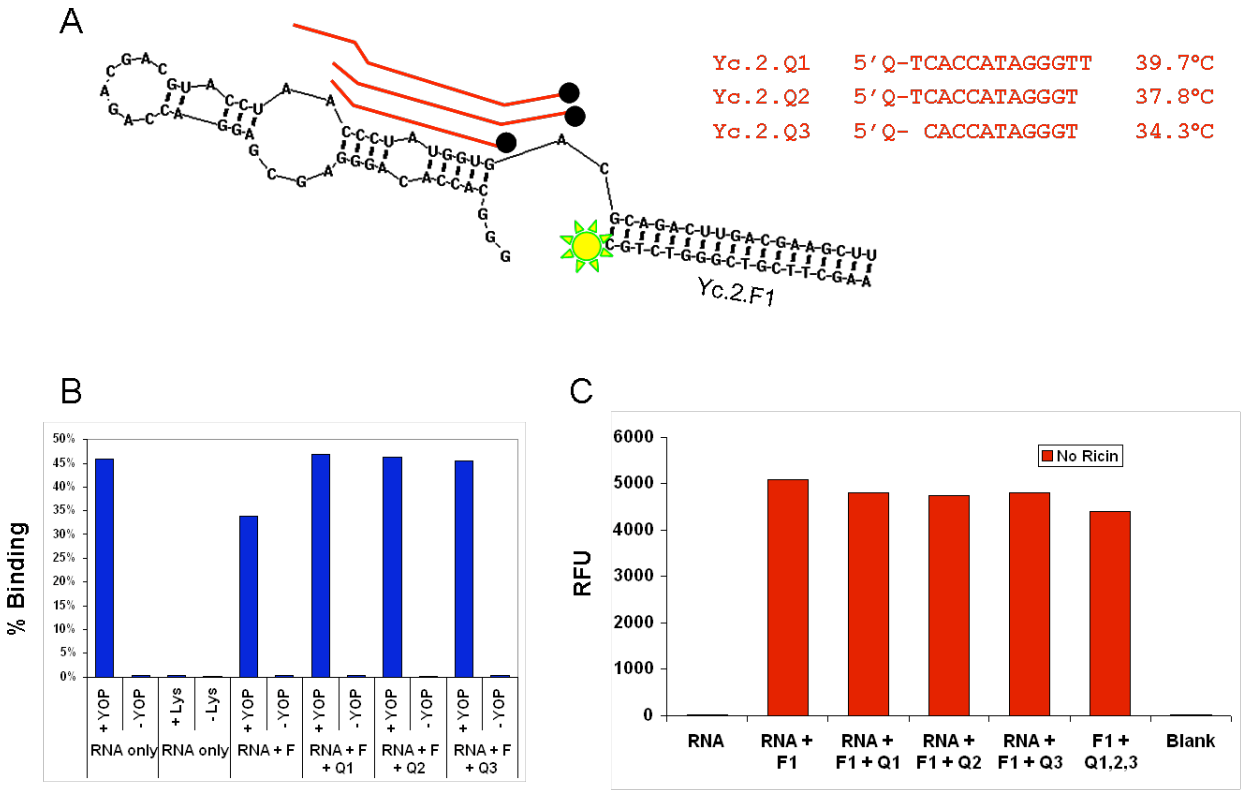


Figure 2.9. Activation and binding assay for anti-Yop aptamer beacon. A.) Anti-YOP aptamer with Yc.2.F1 hybridized and the quencher oligos in red. B.) Binding assay of anti-YOP aptamers N71yc2.d2. C.) Quenching assay of the anti-YOP aptamers N71yc2.d2 with various quenchers after 30 minute equilibration at room temperature.

We therefore tested ligand dependent antisense quencher displacement, and to our surprise, none of the quencher oligonucleotides were capable of hybridizing to the aptamer in the absence of target (**Figure 2.9C**). This was expected for Yc.2.Q3 and possibly for Yc.2.Q2 based on T_m comparisons with previous ricin constructs, however Yc.2.Q1 should have formed a stable duplex with the aptamer. The most likely explanation is that the quencher oligonucleotides failed to destabilize the aptamer structure.

CONCLUSIONS AND PROSPECTS

This project was primarily driven by a collaboration with TIRF Technologies, however we also utilized it as a chance to rationally design aptamer beacons and to decipher to what extent design parameters had on binding or activation. Due to the late arrival of funding, we were able to initiate but not complete the development of protein biosensors. However, we have done exactly what might have been expected within the short timeframe that was allowed: develop initial constructs that are active in binding and can hybridize fluorescently labeled oligonucleotides. One of the advantages of the methods developed here is that since both the fluorescent reporter and the quencher are on separate oligonucleotides, they can be modularly added to the aptamer in various combinations without the need for resynthesis. This design decision saved us a great deal of time during execution.

All that remains is to optimize both the aptamer structure and the quencher oligonucleotide sequence to more finely poise the constructs to signal in the presence of protein analyte. We hypothesized that this poising would have been greatly assisted by the development of computational algorithms to better determine the relationship between design and activity. Nonetheless, the data we derived from these and other constructs fed into additional studies.

MATERIALS AND METHODS

Oligonucleotides

All oligonucleotides were from IDT (Corraville, IA). 5' Iowa Black or 3' Alexa532 were utilized as the quenching or fluorophore moiety (respectively). RNA aptamers were transcribed using AmpliScribe T7 high yield transcription kits (Epicentre,

Madison WI) from 300pmol dsDNA by first extending a pair of overlapping oligonucleotides with SuperScript II (Invitrogen, Carlsbad CA). The oligonucleotide sequences utilized to transcribe the RNA are in **Table 2.1**, whereas the rest of the sequences are listed throughout the figures. The RNA was purified on an 8% denaturing polyacrylamide gel. Strand concentrations were detected by UV spectroscopy at 260nm on a Nanodrop ND-1000 (Wilmington, DE) using extinction coefficients and stock concentrations were adjusted to 10uM for each oligonucleotide.

Melting temperatures were predicted using OligoAnalyzer available from IDT at: http://www.idtdna.com/analyzer/Applications/OligoAnalyzer/.

Activation and Binding Assays

Activation assays were performed by combining the three oligonucleotides in a 1:2:4 ratio (200nM fluorophore: 400nM aptamer: 1200nM quencher) in signaling buffer (DNA anti-ricin aptamer buffer: 20mM HEPES pH 7.35, 120mM NaCl, 5mM KCl, 1mM $CaCl_2$, 1mM $MgCl_2$; RNA anti-Ricin aptamer buffer: 1X phosphate-buffered saline, 5mM $MgCl_2$; RNA anti-Yop51 aptamer buffer: 20 mM Tris pH 7.6, 150mM NaCl, 5mM $MgCl_2$, 1mM dithiothreitol). This ratio was chosen in order to reduce background signal that may have been caused by free fluorescently-labeled oligonucleotides. The mixture was heat-denatured at 65°C for 3 minutes, then equilibrated at room temperature for 10 minutes. Assasys were carried out in Corning 96 well ½ area black opaque plates (Corning, NY). Maximum fluorescence was measured in the absence of quencher oligo for each assay.

Aptamers were assayed for activation in the presence or absence of ligand (1.33uM ricin A-chain: Sigma, St. Louis, MO or 1uM tyrosine phosphatase Yop51*: New England Biolabs, Beverly, MA) on a BioTek Synergy HT fluorescent plate reader (Winooski, VT) using the filter sets Ex 516/Em 575 and room temperatures between

25.7° C and 26.8° C. The gain was automatically set in each assay such that a buffer blank yielded a signal of 26 RFU. Activation reaction was monitored at ten minute increments for one hour.

Binding assays were performed as previously published (Hesselberth, Miller et al. 2000). Briefly, 200nM aptamer or aptamer plus equimolar concentration of protein was incubated in 100uL reaction with 1X signaling buffer for 30 minutes. Binding and non-binding aptamers were separated on a 0.45um HAWP filter (Millipore, Bedford, MA) and washed with 3 volumes of signaling buffer. Non-binding aptamers were captured on nylon filters (Amersham Biosciences, Piscataway, NJ). When fluorescent or quencher oligonucleotides were utilizedd, the ratio was consistent with activation assays. Percent binding was measured by subtracting background filter signal and dividing signal present on the HAWP filter (bound RNA) with combined signal from both the HAWP and nylon filters (total RNA).

Name	Sequence (5' --> 3')
RA80.d5f	GGATCCACATCTACGAATTCAGGGGACGTAGCAATGACTGAGATGCTGGGTTC
RA80.d5r	GAACCCAGCATCTCAGTCATTGCTACGTCCCCTGAATTCGTAGATGTGGATCC
RA80.d6f	CCACATCTACGAATTCAGGGGACGTAGCAATGACTGAGATGCTGG
RA80.d6r	CCAGCATCTCAGTCATTGCTACGTCCCCTGAATTCGTAGATGTGG
RA80.d7f	CGTCTACGAATTCAGGGGACGTAGCAATGACTGAGACG
RA80.d7r	CGTCTCAGTCATTGCTACGTCCCCTGAATTCGTAGACG
RA80.1.d9F	TAATACGACTCACTATTAGGCCACATCTACGAATTCAGGGGACGTAGC
RA80.1.d9cR	AAAAACCAGCATCTCAGTCATTGCTACGTCCCCTGAATTCG
RA80.1.d10F	TAATACGACTCACTATTAGGCCACAUCUACGAAUUCAGGGGACGUAG
RA80.1.d10cR	CCAGCATCTCAGTCATTGCTACGTCCCCTGAATTCG
RA80.1.d11F	TAATACGACTCACTATAGGGAATGGATCCACATCTACGAATTCAGGGGACGTAGCAATG
RA80.1.d11F.2	TAATACGACTCACTATAGGGAAGATCCACATCTACGAATTCAGGGGACGTAGCAATG
RA80.1.d11F.4	TAATACGACTCACTATAGGGAATCCACATCTACGAATTCAGGGGACGTAGCAATG
RA80.1.d11F.6	TAATACGACTCACTATAGGGAACACATCTACGAATTCAGGGGACGTAGCAATG
RA80.1.d11F.8	TAATACGACTCACTATAGGGAACATCTACGAATTCAGGGGACGTAGCAATG
RA80.1.d11Rc	CCAAGCTTCGTCAAGTCTGCAGTGAACCCAGCATCTCAGTCATTGCTACGTCCCCTGA
N71yc2.d2.F	TAATACGACTCACTATAGGGCACCACAGGGAGCGAGGACCAGACGACGTACCTAACC
N71yc2.d2.Rc	AAGCTTCGTCAAGTCTGCGTCACCATAGGGTTAGGTACGTCGTCTGGTCCT

Table 2.1. Sequences utilized to transcribe RNA for tripartite aptamer beacons.

Chapter 3: Computational selection of nucleic acid biosensors via a slip structure model

INTRODUCTION

Aptamers and ribozymes can be adapted to serve as biosensors by a variety of methods (Breaker 2002; Silverman 2003; Verma, Jager et al. 2003; Nutiu and Li 2005). Most of these methods rely upon the aptamer or ribozyme undergoing a ligand-dependent conformational change. A number of authors have generated aptamer beacons in which a non-native conformation is established by base-pairing, and the cognate ligand then stabilizes the native conformation (Stojanovic, de Prada et al. 2001; Fang, Sen et al. 2003). Ligand-dependent changes between non-native and native conformations can be read out by appending reporter molecules to the aptamer; for example, some aptamer beacons are hairpin stem structures in which a fluorophore and quencher are aligned in the non-native conformation, and split apart upon ligand-binding and conformational rearrangement, leading to fluorescence (Hamaguchi, Ellington et al. 2001).

Ligand-dependent conformational changes can also be introduced into ribozymes by design or by selection, yielding allosteric ribozymes (so-called aptazymes). For example, an aptamer that undergoes a flavin mononucleotide (FMN) dependent conformational change was appended to a hammerhead ribozyme; ligand-binding yielded a coupled change in the conformation of the active site and a concomitant increase or decrease in catalytic activity (Soukup and Breaker 1999). Selections for RNA or DNA catalysis can also be designed in which the only successful catalysts will be those that can be activated by or utilize an introduced effector molecule (Roth and Breaker 2004).

To date, though, almost all of these aptamer and aptazyme biosensors have been generated by empirical design (Tang and Breaker 1997) or *in vitro* selection (Soukup and Breaker 1999; Robertson, Knudsen et al. 2004; Nutiu and Li 2005). However, algorithms for the prediction of nucleic acid secondary structure have advanced to the point where nucleic acid secondary structures can be rapidly enumerated based on nucleic acid sequence (Hofacker, Fontana et al. 1994; Zuker 2000). Likewise, probabilistic methods have been developed to sample populations of possible secondary structures and to predict structural stability (McCaskill 1990; Ding and Lawrence 2003). There have also been theoretical treatments of what computational design criteria will maximize sensitivity and specificity in nucleic acid sensors (Dirks, Lin et al. 2004). Therefore, it should be possible to develop algorithms for predicting the sequences of aptamers or aptazymes that will undergo ligand-dependent conformational changes and therefore function effectively as biosensors.

The Breaker lab has developed a computational method for developing allosteric ribozymes activated by nucleic acid sequences (Penchovsky and Breaker 2005). In their work, a virtual pool of RNA molecules was created that contained a randomized region inserted into the middle of the hammerhead ribozyme. An algorithm was then used to predict the dominant secondary structures in the presence and absence of an oligonucleotide effector that was complementary to a given sequence in the randomized region. For example, using this algorithm it proved possible to predict the sequence of ribozymes that would be inactive in the absence of an oligonucleotide effector (due to internal pairings that disrupted the ribozymes), and then would undergo structural rearrangement in the presence of the oligonucleotide effector returning the ribozymes to their native structures. However, computational methods have not been applied to the

more difficult problem of controlling nucleic acid conformational change by non-nucleic acid effectors.

In order to develop prediction methods for nucleic acid biosensors activated by analytes other than oligonucleotides, we wished to first explore a conformational transition that was reasonably well understood and that was computationally tractable. Breaker and co-workers had originally selected flavin-dependent hammerhead aptazymes from pools in which the joining region between an anti-flavin aptamer and the core catalytic domain of the hammerhead ribozyme were randomized (Soukup and Breaker 1999) At least one class of flavin-dependent hammerhead aptazymes was hypothesized to undergo a 'slip structure' conformational change in which the two strands of the joining region between the aptamer and the ribozyme realigned themselves ('changed register') upon the addition of flavin (**Figure 3.1A**).

Herein, we describe the development of an algorithm for assessing the slip structure model, and the application of this algorithm to both designed aptazymes and designed aptamer beacons. Computational automation of the algorithm should allow the structures and minimum free energies of millions of different sequences to be computed in hours on a typical desktop computer. We have validated the utility of the algorithm by experimentally assaying a number of predicted constructs. In general, the model performs as well or better than empirical design, identifying aptazymes that are activated by up to 60-fold by small organic ligands and aptamer beacons that are activated by up to 10-fold by a protein ligand.

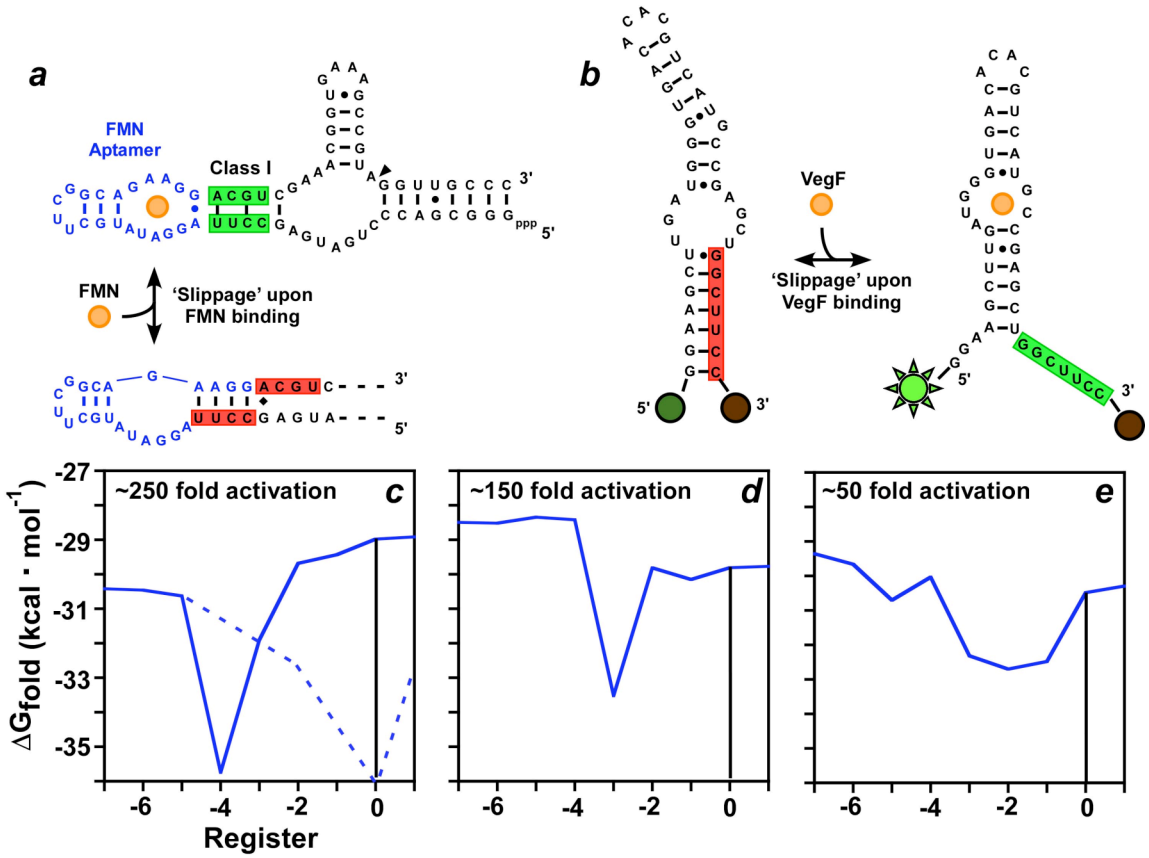


Figure 3.1. Energetic profile model of a slip structure. **a.b.** Schematic of 'slip structure' activation. In the absence of ligand (FMN or VegF), the joining region is trapped in an inactive conformation (red). A structural rearrangement occurs upon ligand binding, shifting the joining region to the 'active register' (green) and potentiating either catalysis (cleavage site denoted by filled arrow in **a.**) or binding. **c,d,e.** Energetic profiles of three previously selected, FMN-responsive hammerhead aptazymes (Class I, Class III, and Class V respectively). The free energies of folding (ΔG_{fold}) in the absence of ligand are plotted versus the register of the joining region. Position '0' (vertical line) denotes the 'active' register of the ribozyme. The dashed line in **c** is a hypothetical energetic profile following ligand-binding.

RESULTS AND DISCUSSION

Development and computational assessment of the slip structure model

Aptazymes have been designed and selected in which ligand-binding aptamer domains are connected to the catalytic domain via a stem or 'joining region.' Breaker and co-workers have previously postulated that for some aptazymes the mechanism of activation involves a 'slip structure' in which ligand-binding induces a local reorientation of a stem and a corresponding change in ribozyme activity. The slip structure model tacitly assumes that there are at least two conformational states for an aptazyme: an unliganded, inactive conformer in which the joining region was 'out-of-register,' and a liganded, active conformer in which the joining region was 'in register.' In this model, the actual conformational change between the two structures could involve sequential changes in register: the stepwise movement of one strand of a joining region relative to the other (**Figure 3.1A**).

In order to assess whether there might be one or more stable, out-of-register conformers for previously selected FMN-aptazymes, we constructed a plot of the Gibbs energy of folding (ΔG_{fold}) for each possible base-pair register within their joining regions (**Figure 3.1C, D, E**). An 'active register' (denoted '0') was first defined based on the known secondary structure of the hammerhead ribozyme. All conformations near the active register were then computationally generated by simply shifting the lower (5'-most) strand of the joining region relative to the upper (3'-most) strand, and determining the predicted folding energy at each register (see also **Figure 3.6**). Conceptually, the energetic profiles that were generated describe the changes in free energy that the aptazymes would undergo as they 'slipped' from one register to the next.

The free energy profiles generated for the different flavin-activated hammerhead aptazymes are consistent with activation via a slip structure mechanism. In the

previously selected Class I aptazyme one strand of the joining region (CCUU) is predicted to base-pair with a complementary portion of the anti-FMN aptamer (**Figure 3.1A**). The assumption of this 'inactive register' effectively disrupts the structure of both the hammerhead and the anti-FMN aptamer (Fan, Suri et al. 1996). The free energy of flavin-binding would stabilize the formation of a different stem structure containing only two Watson-Crick base-pairs in the joining region. In this 'active register' aptamer residues important for flavin-binding (such as the terminal A-G pairing, Fan, Suri et al. 1996) would no longer be sequestered and the hammerhead ribozyme could in turn assume an active conformation. The energetic profiles of two other flavin-activated hammerhead aptazymes with different joining regions (Class III and Class V) were also constructed. As was the case with the Class I aptazyme, out-of-active-register free energy minima were observed (**Figures 3.1D, E**). The free energy of FMN binding to the aptamer (-8.6 kcal/mol) is more than enough to overcome the predicted energetic differences between the active and inactive registers.

The prediction that there are active and inactive registers is consistent with the available structural data. A hammerhead aptazyme that contained an anti-FMN aptamer and a Class I communication module was probed using an 'in-line assay' that exploited conformation-dependent differences in hydrolysis (Soukup and Breaker 1999; Jose, Soukup et al. 2001). Flexible residues, such as those not ensconced in helices, show greater cleavage activity than residues that retain a fixed conformation. The residues that we predict to be unpaired in the unliganded, inactive register were in fact cleaved (see figure 4 in Jose, Soukup et al. 2001). Conversely, these residues should assume a more stable conformation in the liganded, active register, and the cleavage patterns were found to be correspondingly suppressed.

The predicted free energy profiles are not only consistent with activation via a slip structure mechanism, but could also account for the relative levels of activation. The Class I aptazyme was predicted to have the greatest energetic separation between the stable active and inactive registers at nearly 7 kcal/mol (**Figure 3.1C**), and also showed the greatest effector-activation. Class III and Class V have smaller predicted energy barriers of 4 kcal/mol and 2.5 kcal/mol respectively, and exhibited correspondingly smaller activations. Both the height of the energy barrier and the distance between the stable inactive register and the less stable active register may affect the level of aptazyme activation.

To further validate the utility of using energetic descriptions for slip structure mechanisms, energetic profiles for previously selected hammerhead aptazymes (Soukup and Breaker 1999) that were inhibited rather than activated by FMN were constructed (data not shown). As expected, there were energetic minima at the active registers and less stable structures in adjacent registers; in other words, the predicted energy profiles were the converse of the energetic profiles predicted for flavin-activated aptazymes. Presumably, FMN-binding leads to a reorganization of the joining region such that an inactive register becomes favored relative to the active register.

Computational selection of allosteric ribozymes

To the extent that the slip structure model can be used to explain the activation parameters of known hammerhead aptazymes, it should prove possible to use this model to predict the sequences of new hammerhead aptazymes. To this end, we developed a simple algorithm to select slip structure aptazymes from computationally-generated random sequence populations. Energetic profiles similar to those shown in **Figure 3.1** were computed for each member of a pool in which the hammerhead ribozyme was joined to the anti-FMN aptamer via a symmetric N3-N3 joining region (4^6 = 4,096

variants; **Figure 3.2A**). The energetic profiles of each candidate in the pool were compared to a reference energetic profile of a known hammerhead aptazyme (Class VI; **Figure 3.2A**, blue line). Aptazyme variants that had similar energy profiles to the Class VI hammerhead aptazyme were selected for further analysis. A total of 23 candidate aptazyme sequences were chosen that most closely matched the energy profile Class VI aptazyme; the energetic profiles of two of these aptazymes are shown in **Figure 3.2A**. These energy profiles differed substantially from the norm. The black line in **Figure 3.2A** corresponds to the average energetic composition of the un-selected pool.

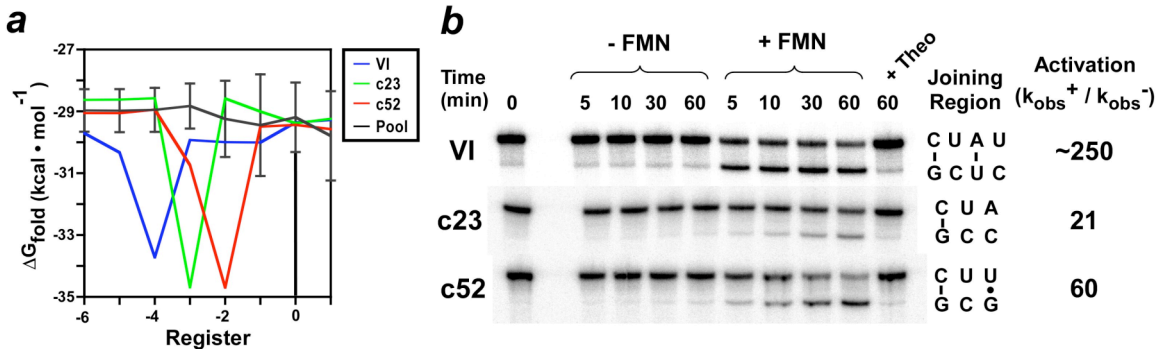


Figure 3.2. Computationally selected hammerhead aptazyme variants. **a.** The energetic profiles of every member of a pool in which an anti-flavin aptamer was joined to the hammerhead ribozyme via a symmetric, N3-N3 stem were compared to a reference energy profile (Class VI hammerhead aptazyme, blue line). The average energetic profile and distribution (error bars) of the pool are shown in black. The energetic profiles of two candidate aptazymes are shown in green (aptazyme c23) and red (aptazyme c52). **b.** Activity assays with selected variants. Cleavage as a function of time following the addition of flavin was analyzed by polyacrylamide gel electrophoresis. Non-specific activation by theophylline (Theo) was also monitored. The sequences of the joining regions for the aptazymes and the calculated levels of activation in the presence of FMN are listed to the right.

All 23 designed aptazymes were synthesized and the aptazyme variants were assayed *in vitro*. Since these aptazymes were predicted to structurally rearrange from an inactive to an active conformation only in the presence of effector, they should show substantially greater cleavage activity only in the presence of FMN. Three variants

exhibited greater than 4-fold activation in the presence of FMN (**Table 3.1**), and two of these, c23 and c52, exhibited activations (k_{obs}^{+} / k_{obs}^{-}) of 21- and 60-fold, respectively (**Figure 3.2B**). Of the remaining aptazymes eight were constitutively inactive and twelve were constitutively active.

A control was carried out in which 23 aptazyme variants were selected at random from the population. Three were slightly activated by FMN, although the best variant was activated only 5-fold (**Table 3.1**), as compared with over 20-fold for the computationally selected aptazymes. In addition, five of the randomly selected aptazyme variants were found to be inactivated by FMN, whereas no computationally selected aptazymes were inactivated.

Pool	Ribozyme	Aptamer	State	Register[a]	Off[b]	On[c]	Activated	Inhibited	Identified[d]	Tested[e]
N3-N3	HH	FMN	Random	n/a	3	12	3	5	-	23
N3-N3	HH	FMN	Selected	-4	8	12	3	0	23	23
N3-N3	HH	Theo	Selected (1)	-4	0	3	1	0	4	4
N3-N3	HH	Theo	Selected (2)	-8	4	19	1	0	44	24
N4-N4	HH	Theo	Random	n/a	3	7	0	0	-	10
N4-N4	HH	Theo	Selected	-8	17	8	2	0	48	27
N5-N5	HH	Theo	Random	n/a	4	6	0	0	-	10
N5-N5	HH	Theo	Selected	-8	23	1	1	0	11,637 (68)	25
N3-N3	Hairpin	FMN	Selected	-4	12	0	1	0	13	13
N5-N5	HH	FMN	Selected	n/a	12	1	2	0	8,607	15

Table 3.1. Summary of computationally identified and experimentally tested aptazymes.
HH = Hammerhead
[a] 'Register' indicates the required distance between the stable active and inactive registers. This distance could vary up to 2 positions in the algorithm, as stated in **Figure 3.2.**
[b] 'Off'indicates ribozymes that were found to be constitutively inactive.
[c] 'On' indicates ribozymes that were found to be constitutively active.
[d] 'Identified' gives the number of ribozymes picked by the algorithm.
[e] 'Tested' gives the number of ribozymes synthesized and assayed.

Varying the aptamer sensor domain

Some joining regions identified by *in vitro* selection experiments have been shown to act as 'communication modules' that can lead to the activation of ribozymes by many different aptamers (Soukup and Breaker 1999). We therefore replaced the anti-flavin aptamer on the two best computationally selected aptazymes with an anti-theophylline aptamer which should in turn alter the aptazyme's effector specificity from FMN to theophylline. One of the computationally-selected joining regions (c23) conferred roughly the same activation by theophylline (17-fold); the other joining region (c52) showed no activation.

Given this success and the fact that theophylline-activated hammerhead aptazymes had previously been selected (Jenison, Gill et al. 1994; Soukup, Emilsson et al. 2000), the anti-theophylline aptamer was computationally appended to a N3-N3 pool, and the pool was sieved using energy parameters identical to those that had been used to identify flavin-dependent aptazymes. Four other sequences were identified as candidate aptazymes, and were assayed for theophylline responsivity (**Table 3.1**). A single variant from the four exhibited a 14-fold increase in rate upon the addition of theophylline.

Varying the length of the joining region

Since computational selections with short joining regions had proved successful, we were interested whether longer, more complex joining regions could also be sieved. The anti-theophylline aptamer was computationally joined to the hammerhead ribozyme by both N4-N4 or N5-N5 stems and the resultant pools were sieved using the slip structure energy model. However, the distance between the active and inactive registers was changed from 2-4 to 4-8, reflecting the best-activated variants from the experimental N5-N5 selection (**Table 3.1**). From the pool with a N4-N4 joining region 48 candidates were identified and 27 were experimentally assayed for self-cleavage in the presence of

theophylline. However, from the pool containing a N5-N5 joining region, 11,637 sequences (representing 1.1% of the original pool) were identified, despite the fact that the same parameter sets were utilized. The complexity of this population was reduced by taking only one variant from selected sets whose 5'-most N5 regions were identical. This left 68 candidate sequences, and 25 were experimentally assayed. Two variants from the N4-N4 pool were found to be activated by theophylline (9- and 20-fold), and one clone from the N5-N5 pool was theophylline-responsive (16-fold).

To again ensure that we had not identified aptazymes by chance alone, ten designs from the N4-N4 and N5-N5 pools were picked at random from these pools, synthesized, and assayed for activation in the presence of theophylline. Not only were no effector-dependent variants found, but the distribution of active and inactive variants was very different between the randomly selected and computationally sieved sequences (**Table 3.1**). The randomly selected sequences favored constitutively active variants, while the computationally sieved sequences contained many fewer constitutively active variants than inactive variants.

Varying the ribozyme response domain

While various aptamer domains could be easily interchanged using the slip structure model, we also wanted to test whether the same model could be generalized to other catalytic platforms. The hairpin ribozyme was chosen as a candidate for engineering because like the hammerhead its structure also consisted of stacked helices (Rupert and Ferre-D'Amare 2001).

The anti-FMN aptamer was appended to the end of loop B of the hairpin ribozyme via a symmetric N3-N3 random region, and 13 candidate aptazymes were identified using the same slip structure energy model that had proven successful for hammerhead aptazymes. All 13 sequences were then assayed *in vitro* for their ability to

undergo cleavage in the presence of FMN. A single sequence was found to be activated 10-fold in the presence of FMN (**Figure 3.3**).

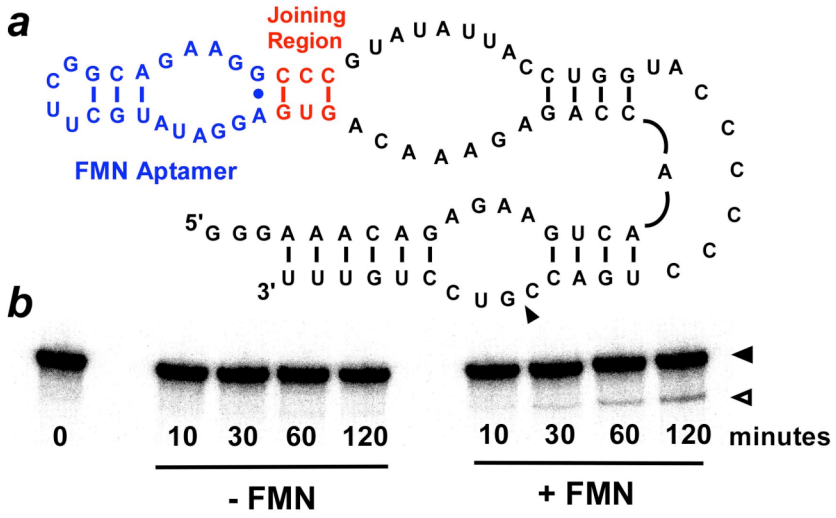


Figure 3.3. Computationally selected hairpin aptazyme variants. **a**. The hairpin ribozyme (black, cleavage site marked with filled arrow) was modified with a randomized stem region (red) and an anti-FMN aptamer (blue). Selection identified an aptazyme whose joining region facilitated activation in the presence of FMN. **b**. Activity assay with the selected variant. The filled and open arrows mark the positions of uncleaved and cleaved products, respectively.

Computational selection of signaling aptamers

Following up on our success using the slip structure energy model to computationally predict aptazyme sequences, we sought to use the model to generate conformation-switching aptamers that could be used as optical biosensors. A non-binding or inactive register was generated by manipulating the terminal stem of a hairpin-like aptamer. In the model, the addition of ligand should shift the register of the aptamer stem from the non-binding conformation to the original binding conformation. Instead of detecting ligand-activated ribozyme cleavage, ligand-binding was coupled to structural changes that induced fluorescent signaling, much like a molecular beacon (**Figure 3.1B**).

We also attempted to extend the range of analytes that could activate computationally predicted biosensors from small organic molecules (such as theophylline and FMN) to proteins (such as vascular endothelial growth factor (VegF)). A series of random sequence libraries of 4 to 9 residues in length (349,440 variants total) were created *in silico* and appended to the 3' end of an anti-VegF aptamer 44T (Jellinek, Green et al. 1994). We chose this aptamer because of the structural similarities to the FMN aptamer including a stem loop and internal bulge. Minimum free energies were then calculated for each register of each member of the library.

Since no VegF-activated aptamer beacon existed, no reference energy profile was available to sieve putative sequences from the library. Therefore, we devised a set of criteria that could be sequentially applied (see also **Figure 3.7**): (i) We first rank-ordered the sequences based on how well-defined the local minima were at the inactive register. Those aptamer beacons in which the free energy of folding at the inactive register differed most from the free energies of folding for the registers to either side were ranked most highly (see also **Figure 3.4B**). By choosing well-defined local minima we hoped to reduce folding transitions in the absence of ligand and hence background signal (Flamm, Hofacker et al. 2001). (ii) We then chose an energetic difference between the active and inactive registers that we hoped would lead to an increase in signal in the presence of ligand. An arbitrary cutoff value of at least 3.5 kcal mol^{-1} energy difference between the active and inactive registers ($\Delta G_{fold,\ inactive} - \Delta G_{fold,\ active}$) was initially chosen. This arbitrary choice can be further refined by experimental feedback, as described below.

a

Name	Sequence
44t.mab1	GGAAGCUUGAUGGGUGACACACGUCAUGCCGAGCUGGCUUCC
44t.mab2	GGAAGCUUGAUGGGUGACACACGUCAUGCCGAGCUAGCUUCC
44t.mab3	GGAAGCUUGAUGGGUGACACACGUCAUGCCGAGCUGGCUUCU
44t.mab4	GGAAGCUUGAUGGGUGACACACGUCAUGCCGAGCUAGCUUCU
44t.N7A	GGAAGCUUGAUGGGUGACACACGUCAUGCCGAGCUAAAAAAA
Rand.mab1	GGAAGUAACUCGAAACUACCUACCAUGCUGUCGCCGGCUUCC

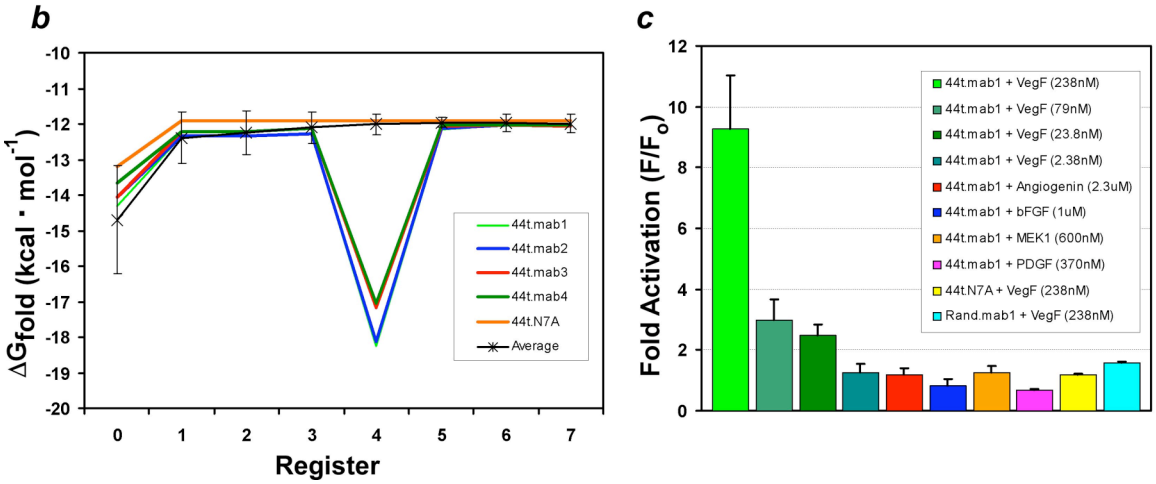


Figure 3.4. Design of a VegF-responsive aptamer beacon. **a.** Following computational selection (as described in the text), four variants (44t.mab1-mab4) were identified from a N7 pool that had an energy minima at register 4. The sequences of these variants are shown, along with two negative controls (44t.N7A and Rand.mab1). The aptamer beacon 44t.mab1 was used in experimental studies. The randomized region of 44t.mab1 is shown in red, and residues in other aptamer beacons that differ from 44t.mab1 are shown in blue. **b.** Energetic profiles for selected aptamer beacons. In addition to the energetic profiles for 44t.mab1-mab4 and 44t.N7A, the average energetic profile of all 16,384 variants is shown in black with error bars representing the standard deviation. **c.** Sensitivity and selectivity assays of the designed aptamer beacon. The aptamer beacon 44t.mba1 (2 μM) was assayed at four VegF concentrations and with 4 other proteins involved in angiogenesis (concentrations shown). Also, negative controls were assayed with VegF (238 nM). These experiments were performed in 30 μL reactions on an MJ Opticon at 37° C. VegF-responsive aptamer beacons were at 2 μM final concentration and readings were taken after 60 minutes

Based on these criteria, no variants from the N4, N5 or N6 pools were predicted to function as aptamer beacons. A total of 7 sequences were selected from the N7 pool (**Figure 3.4A**), 58 from the N8 pool, and 514 from the N9 pool. The increase in predicted active variants roughly corresponds to the increase in pool complexity. In general, we found that the largest energy differences between the active and inactive registers (criterion (ii), above) did not exceed 7 kcal mol^{-1} ($\Delta G_{fold,\ inactive}$ - $\Delta G_{fold,\ active}$). This energy difference should have been more than adequate for conformational transduction, as the binding energy for VegF has been measured to be -12.7 kcal mol^{-1} (Jellinek et al., 1994). Because of the difficulties inherent in the synthesis of long, derivatized RNA molecules, we wished to choose a single variant that would have the best chance of showing ligand-dependent signaling. Of the seven sequences chosen from the shortest, most synthetically accessible pool (N7), only four could form terminal base-pairings that would place fluorophores and quenchers adjacent to one another. Of these four, 44t.mab1 had the best-defined local minimum (criterion (i), ΔG_{fold} = -5.97 kcal mol^{-1}; **Figure 3.4B**); and had the second largest energetic differences between the active and inactive registers (criterion (ii), $\Delta G_{fold,\ inactive}$ - $\Delta G_{fold,\ active}$ = 3.91 kcal mol^{-1}); and also formed beacon like terminal base pairs in the inactive conformer.

We synthesized 44t.mab1 as an aptamer beacon with fluorescein and DABCYL incorporated at the 5' and 3' ends, respectively. In the absence of VegF, the aptamer should form a structure in which DABCYL quenches the fluorescein, while in the presence of VegF the aptamer should form a structure in which fluorescence is not quenched. The computationally predicted signaling aptamer was activated 10-fold by 238 nM VegF (**Figure 3.4C**), and displayed optimum activity at 37° C (data not shown). These results were especially impressive considering that other experimentalists have observed only 3- to 4-fold activation at similar protein concentrations when examining

similar constructs designed 'by hand' (Nutiu and Li, 2003). The aptamer beacon was both sensitive and selective for VegF. The beacon displayed a two-fold fluorescence increase at analyte concentrations as low a 24 nM (1 μg ml^{-1}), and was not activated by non-cognate growth factors including PDGF, bFGF, or angiogenin (**Figure 3.4C**).

The kinetics of conformational change and activation were, however, relatively slow: it took 15 minutes to reach 50% activation and about 75 minutes to reach full activation (data not shown). The kinetics of activation are consistent with a model in which the rearrangement of the engineered stem precedes protein binding. The observed responsivity is comparable to the typical time required for the development of signals from protein-binding aptamer microarrays and much faster than the time required for antibody microarrays that typically require complicated post-binding procedures (such as the binding of secondary reagents in sandwiches and wash steps (Collett, Cho et al. 2005)). Thus designed signaling aptamers should be compatible with high-throughput, chip-based analyses.

Mutational analysis of the computationally designed signaling aptamer

To better assess how the selection criteria related to experimental performance, we rationally introduced mutations into the VegF signaling aptamer that were predicted to either stabilize or destabilize the inactive conformer (**Figure 3.5A**). As can be seen in **Figure 3.5B**, stabilizing mutations had no affect on activation, suggesting that the criteria used for the selection of the inactive register overpredicted stability. Consistent with this interpretation, the destabilizing mutations increased specific activation from approximately 1.5-fold at low (40 nM) protein concentrations to 3-fold (**Figure 3.5B**). As expected, the destabilizing mutations also increased background fluorescence (pre-formation of the active conformer) prior to protein addition. The initial successes of the slip structure model now provide a basis for further determining how systematically

varying the free energy parameters for computationally-derived biosensors affects their overall performance.

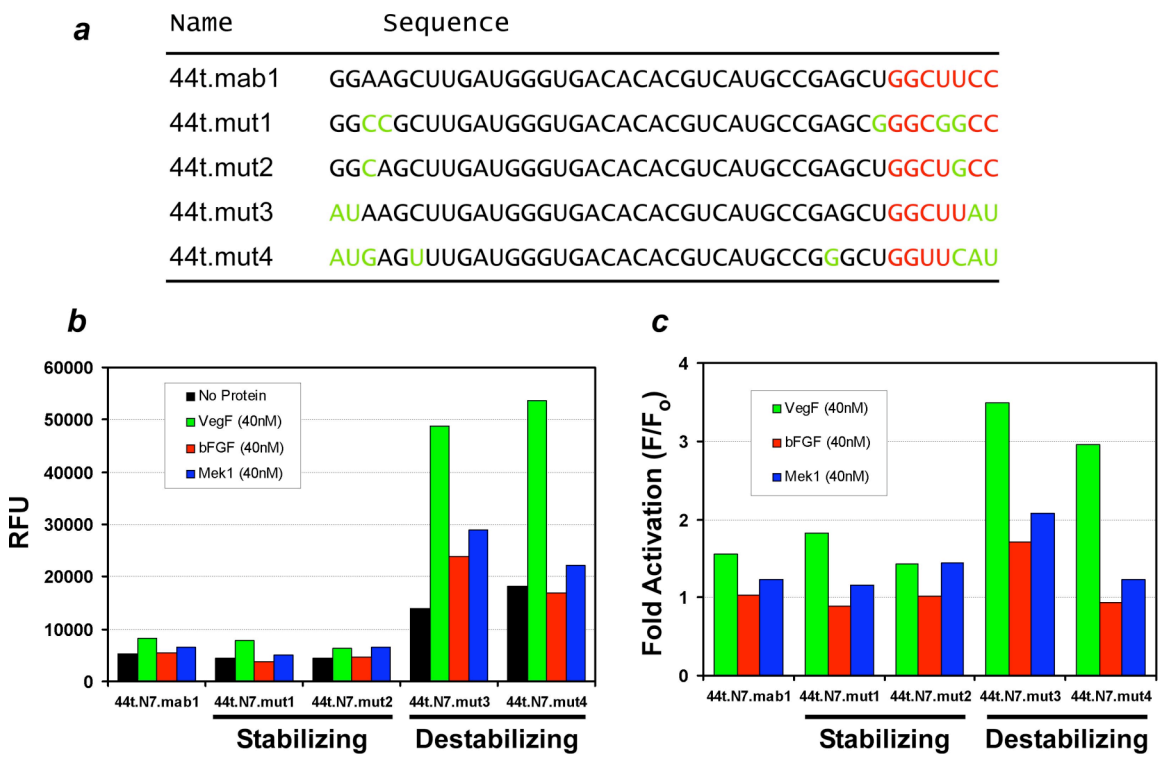


Figure 3.5. Modulating the performance of the VegF-responsive aptamer beacon. **a.** A series of predicted stabilizing (44t.mut1-2) or destabilizing mutations (44t.mut3-4) of 44t.mab1 were designed. The randomized region of 44t.mab1 is red, and mutational changes are green. **b.** Observed fluorescence (RFU) of 44t.mab1 and its mutations with various angiogenesis proteins. **c.** Same results as (**b**), but represented in terms of fold fluorescence increase in the presence of protein. Reactions were carried out in TMK buffer on a Tecan Safire at 40° C. Aptamer beacons were at 200 nM final concentration and readings were taken after 100 minutes.

CONCLUSION

While *in vitro* selection methods are extremely robust, the ability to use computational design methods for the generation of biosensors would be extremely

valuable for a variety of reasons. First, the ability to design nucleic acid biosensors further validates the utility of methods used for the computational prediction of nucleic acid secondary structures. In particular, by showing that a slip structural model accords with both computational predictions and experimental data it may prove possible to use similar models to better understand conformational transitions in both natural and selected functional nucleic acids. Second, computational selection should be inherently faster than experimental selection methods, and thus the time and effort required for the development of biosensors may be greatly decreased. Based on the methods described in this paper, only a few sequences had to be synthesized and assayed to identify functional biosensors, while following selection experiments numerous different sequences typically have to be selected, characterized, and screened for their ability to function as biosensors. Ultimately, the computational sieving algorithm can be employed within hours to quickly generate a set of aptazymes or aptamer beacons corresponding to any of the large number of aptamers that have already been found. Coupled with high-throughput, chip-based nucleic acid synthesis (McGall and Fidanza 2001; Nuwaysir, Huang et al. 2002) it may therefore prove possible to rapidly generate biosensor chips for the detection of multiple analytes. Lastly, to the extent that computational models match experimental results, it should be possible to finely control the design process, generating biosensors with optimal sensitivities, signal-to-noise ratios, and dynamic ranges.

ACKNOWLEDGEMENTS

The authors would like to thank Lauren Ancel Myers for helpful discussions and knowledgeable insights. J.H. and B.H would like to acknowledge partial financial supported from an Integrative Graduate Education and Research Traineeship (IGERT) and the National Science Foundation. This work was funded by a grant from the NIH.

METHODS AND MATERIALS

Computational methods

The free energies of RNA secondary structures were calculated using the ViennaRNA package (Hofacker, Fontana et al. 1994). Scripts to automate the evaluation and comparison of RNA secondary structure profiles were coded in the scripting language Python (http://www.python.org/). These programs first generated nucleic acid sequences containing randomized regions: either a randomized joining region between an aptamer and a ribozyme, or a randomized region at one end of an aptamer. For each sequence in the randomized population, the free energy of folding was calculated along a series of 'registers,' defined as sequential alignments of a randomized region with potential pairing partners in the native portions of an aptazyme or aptamer. Energetic profiles (free energies of folding versus register) were constructed based on these calculations. A flowchart that describes the computational selection of aptazymes is presented in **Figure 3.6**.

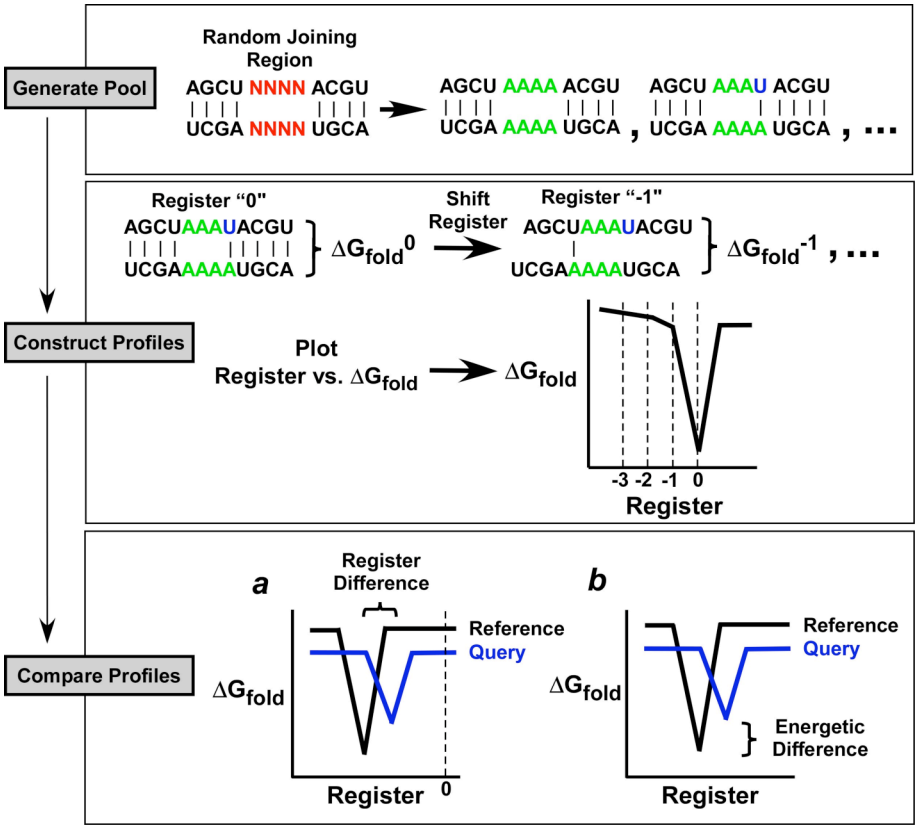


Figure 3.6. Flowchart for computational selection of aptazymes. The sequences of all members of a pool spanning the joining region between aptamer and aptazyme were computationally generated. To construct free energy profiles, the 'register' of one strand in the joining region was fixed relative to the other, and ΔG_{fold} was calculated. The register was then moved by one residue and the ΔG_{fold} was again calculated. The free energy profile (ΔG_{fold} versus register) of each aptazyme was compared with a reference profile generated from an experimentally validated hammerhead aptazyme (Class VI). To compare the energetic profiles of the aptazymes chosen by the algorithm ('Identified,' Table 3.1) two criteria were used: a. the register difference (2-4 in the N3 N3 pool and increased based on pool size), and b. the energetic difference between the minimums. Except for the N5-N5 selection described in the text, relatively few sequences were sieved from computationally generated populations based on these parameters.

The aptazyme variants that best matched a defined profile were selected for further study. For example, the calculated energetic profiles were compared to the energetic profile of an experimentally-validated reference aptazyme. Similarly, aptamer

beacons were rank-ordered based on a number of criteria garnered from the *in silico* aptazyme selection (**Figure 3.7**).

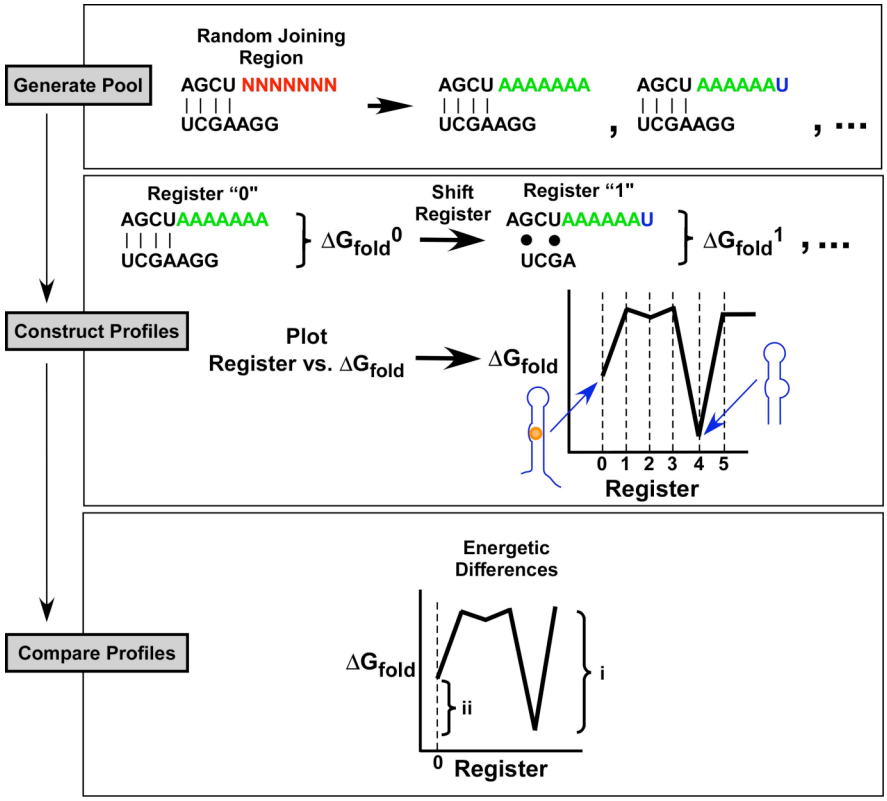


Figure 3.7. Flowchart for computational selection of aptamer beacons. Initially, the sequences of all members of a joining region pool were generated. To construct profiles, the 'register' of the joining region was fixed and the ΔG_{fold} was calculated. The register was then moved as in the aptazyme selection and the ΔG_{fold} was again calculated. A profile of ΔG_{fold} versus register was constructed. Aptamer beacon profiles were ranked and compared based on maximizing the isolation of the inactive register (ΔG_{fold} of the inactive register relative to the registers to either side of it, **i**, in the Figure), and maximizing the energetic difference between the active and inactive registers (**ii**, in the Figure). Except for the N8 and N9 selection, typically relatively few sequences were computationally sieved from the populations based on these parameters.

In vitro transcription of aptazymes

Double-stranded DNA templates used for transcription were made by primer extension of complementary, synthetic DNA oligonucleotides (IDT, Coralville, IA). Radiolabeled RNA was generated by *in vitro* transcription using T7 RNA polymerase with trace amounts of [α-^{32}P] UTP (Perkin-Elmer, Boston, MA), followed by purification on 10% denaturing polyacrylamide gels.

RNA synthesis of aptamer beacons

RNA was synthesized on an Expedite 8909 DNA synthesizer (PE Biosystems, Foster City, CA) using standard phosphoramidite chemistry. 3'-DABCYL (4-(4-dimethylaminophenylazo)benzoic acid) and 5'-Fluorescein (6-FAM) were added using a fluorescein phosphoramidite and a 3'-DABCYL controlled pore glass (CPG) column, respectively. All synthesis reagents were purchased from Glen Research (Sterling, Virginia). Products were purified by reverse phase chromatography or on denaturing polyacrylamide gels. The concentrations of aptamer beacons were determined based on their calculated extinction coefficients.

Aptazyme assays

Radiolabeled RNAs (20 pmoles) were thermally equilibrated by denaturation at 70° C in water followed by cooling to room temperature. Buffer (50 mM Tris-HCl, pH 7.5, 20 mM $MgCl_2$, final concentration) and ligand (FMN at 200 μM, theophylline at 1 mM, or ATP at 1 mM). Reactions (20 μL) were carried out at room temperature, and aliquots were removed at various time points and quenched in stop dye (85% formamide, 0.01% bromophenol blue, 0.01% xylene cyanol and 60 mM EDTA, pH 8.0). Uncleaved

and cleaved ribozymes were separated on 8% denaturing polyacrylamide gels, and radiolabeled bands were quantitated using a PhosphorImager (Molecular Dynamics, Sunnyvale, CA). Rate constants were calculated by plotting the negative natural logarithm of the fraction of uncleaved ribozyme versus time, and determining the slope of a line fitted through at least three time points in the linear range of the assay. Assays were repeated twice or more; observed levels of activation differed by less than two-fold.

Aptamer Beacon assays

Aptamer beacons (50 μM) in buffer (PBS [137 mM NaCl, 2.7 mM KCl, 4.3 mM $Na_2PO_4*7H_2O$, 1.4 mM KH_2PO_4; pH 7.4] or TMK [100 mM Tris pH 7.5; 80 mM KOAc; 10mM $Mg(OAc)_2$]; 30 μL) were denatured at 80° C for three minutes and cooled to room temperature over 5 to 15 minutes. Either an MJ Research DNA Engine Opticon (Waltman, MA) with LED fluorescence excitation at 495 nm and photomultiplier tube (PMT) emission detection at 518 nm or a Tecan Safire (San Jose, CA) fluorescent monochrometer plate reader with excitation at 480/2 nm and emission detection at 520/2 nm were used to assay aptamer beacons. The thermal denaturation profiles for aptamer beacons were determined in the Opticon by obtaining fluorescence readings at temperatures from 25° C to 90° C in 1° C increments and with equilibration at each temperature for 1 min.

In order to measure analyte activation, aptamer beacons were denatured as described above, then allowed to cool to room temperature over 10 minutes. Aptamer beacons were then heated to 30° C in the Safire for 10 minutes, protein was added, and the fluorescence data was taken at two minute intervals for up to two hours. The proteins assayed included recombinant human VegF (R&D Systems, Minneapolis, MN),

recombinant human bFGF (Bachem, King of Prussia, PA), recombinant human PDGF-AB (R&D Systems, Minneapolis, MN), recombinant human angiogenin (Bachem), recombinant human MEK-1 (Santa Cruz Biotechnology, Santa Cruz, CA), glycosylated ricin A chain (Sigma Aldrich, St. Louis, MO), and bovine serum albumin (New England Biolabs, Ipswich, MA). All proteins were initially reconstituted in the buffers described by their manufacturers.

Chapter 4: Computational design and experimental validation of antiswitches.

INTRODUCTION

Antiswitch

One of the current limitations of signaling aptamer beacons is the lack of a signal amplification cascade upon binding. Since aptamer beacons often display reduced binding to their targets compared with the originally selected aptamer (discussed below), detection limits are often reduced. Signal amplification could be utilized to compensate for this limitation. Typically, one aptamer beacon "signal" is related to one binding event and therefore detection is limited to ligand concentration and the sensitivity of the equipment utilized for detection. A relatively recent and wholly unexplored type of signaling aptamer that may overcome this limitation is the assisted refold or antiswitch developed by Travis Bayer and colleagues with the anti-theophylline or anti-tetracycline aptamers (**Figure 1.6A, E**; Bayer and Smolke 2005).

An antiswitch consists of three parts: a hairpin or antiswitch stem, an aptamer and a displacement stem. The switch works by forming a stable conformer that traps or "hides" a sequence in a hairpin stem (**Figure 1.6A** in red) in the absence of ligand. Upon the addition of ligand, the aptamer portion undergoes a conformational change, forming an active binding conformer in which the displacement stem (**Figure 1.6A** in blue) forces the hairpin into a single stranded state. This free single stranded region is capable of performing a number of different tasks, including activating or inhibiting gene expression *in vivo*, as originally published. It is also free to prime PCR or rolling circle amplification (**Figure 1.6C**; Yang, Fung et al. 2007). Although untested, it could initiate a hybridization chain reaction originally developed by Dirks *et al* (Figure 1.6B; 2004).

We provide evidence of antiswitch conformational changes activating molecular beacons in this work (**Figure 1.6D**). Additionally, the antiswitch could provide allostery for oligonucleotide dependent ribozymes (**Figure 1.6F**; Penchovsky and Breaker 2005), be utilized as fuel for DNA machines (Green, Lubrich et al. 2006) or form "toeholds" typically utilized in nucleic acid logic circuits (Kim, White et al. 2006; Seelig, Soloveichik et al. 2006). Researchers have even begun to reversibly modulate enzyme activity through switchable aptamers and single stranded "removal strands" (Friedrichs and Simmel 2007). The possibilities for ligand inducible regulation of single stranded DNA are vast, including application involving oligonucleotide activators.

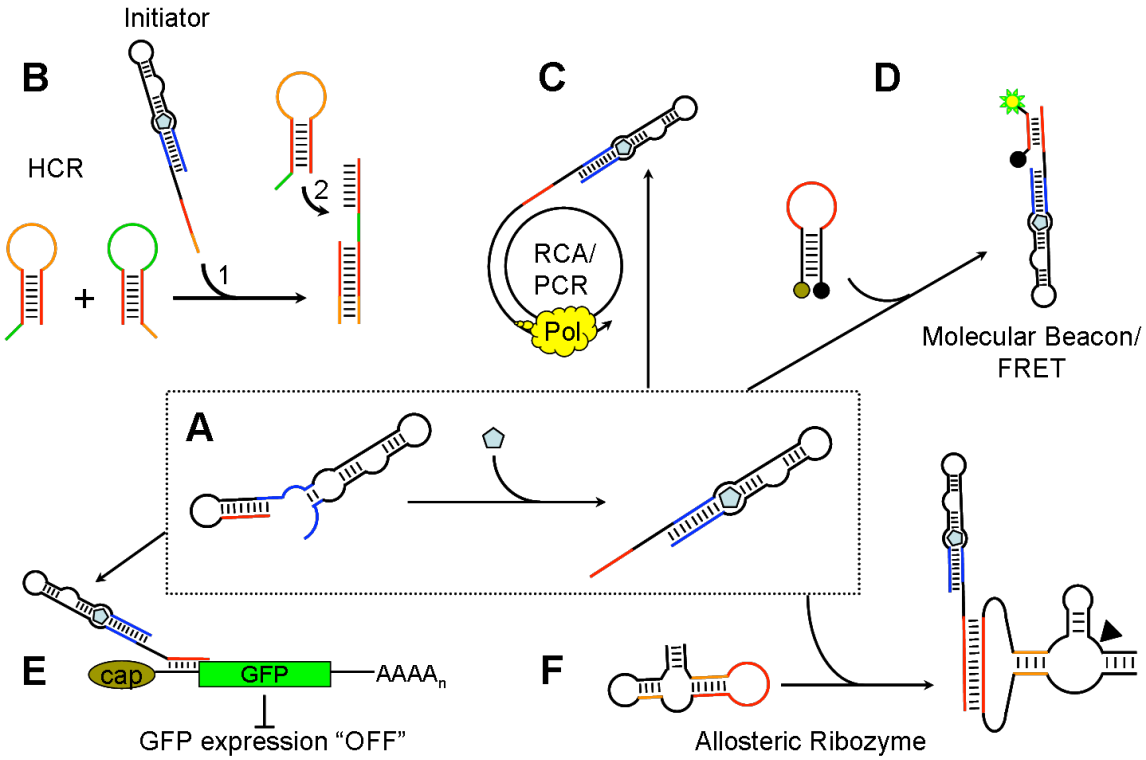


Figure 4.1. Antiswitch based biosensor design. A.) A conformational switch is induced by target (pentagon) binding, in which a sequence (blue) displaces a portion of the antiswitch hairpin (in red) making it available for various signaling pathways (B-F). These pathways include B.) hybridization chain reaction, C.) rolling-circle amplification, D.) molecular beacon activation, E.) modulation of GFP expression *in vivo* and F.) ribozyme activity (cleavage denoted with carrot).

Theophylline Aptamer

In the seminal work by Bayer *et al*, the well characterized anti-theophylline aptamer mTCT8-4 originally selected by Jenison and colleagues (1994) was utilized to activate or inhibit GFP or YFP production. The affinity for theophylline (K_d = 0.29uM) is dependent on divalent metal ions including Mg^{2+}, Mn^{2+} and Co^{2+} (Zimmermann, Wick et al. 2000) that help stabilize the RNA secondary structure (Jenison, Gill et al. 1994). In addition, structural studies have been performed including computational three dimensional modeling (Tung, Oprea et al. 1996) and NMR techniques (Zimmermann, Jenison et al. 1997) that suggest a number of hydrogen bonds, Hoogsteen pairs and noncanonical base triplets help stabilize the tertiary structure. In fact, theophylline interacts to form a base triplet with hydrogen bonds to both C22, U24 found in the upper internal loop (**Figure 4.2B**). Stopped-flow fluorescent spectroscopy analyses have determined that the high binding affinity is due to a stabilization of the binding loop which significantly decreases the dissociation rate (Jucker, Phillips et al. 2003).

Specificity, or the discrimination of closely related analytes is a hallmark of some aptamers. For instance, the anti-L-arginine RNA aptamer displays 12,000 fold discrimination between L-arginine and its enantiomer D-arginine (Geiger, Burgstaller et al. 1996). Likewise, there is a 10,000 fold difference in affinity between theophylline and caffeine by the anti-theophylline aptamer. These two molecules differ only by a methyl group at the N7 position that is responsible for destabilizing the divalent metal binding loop of this aptamer (**Figure 4.2**; Zimmermann, Wick et al. 2000)..

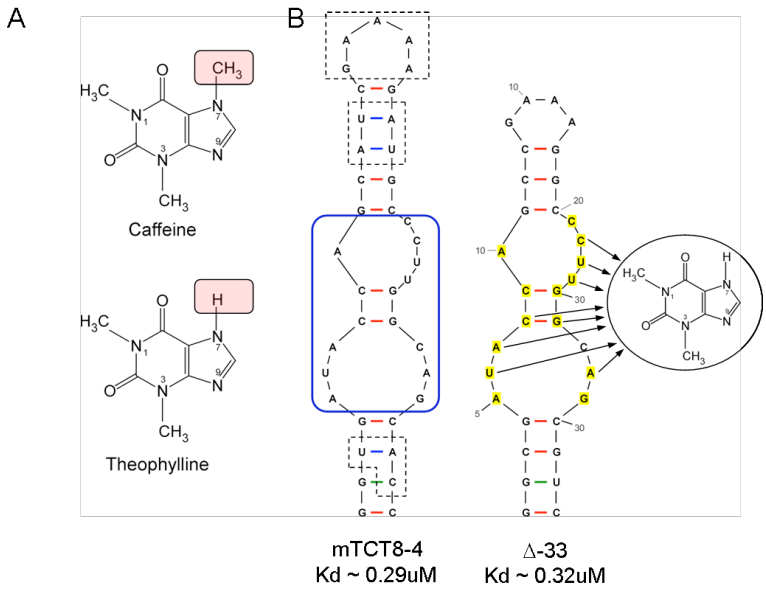


Figure 4.2. Theophylline aptamer and derivatives. A.) Chemical structures of caffeine and theophylline differing in the N7 methyl group (red box). B.) Secondary structure of two aptamers that display high binding affinities for theophylline. The blue box is the conserved aptamer sequence across a number of aptamer clones originally selected (Jenison et al). The dashed boxes represent sequence differences between mTCT-8-4 and Δ-33. The yellow colored bases are involved in non-covalent bonds with either theophylline (arrows) or represent noncanonical base pairs necessary for binding theophylline. C32 is nonspecific and can be an A. Base numbering is specific to Δ-33. Adapted from Zimmerman et al (1997).

Theophylline biosensors

With rational design it has proven possible to modulate various processes through the incorporation of a theophylline aptamer. For instance, the aptamer was engineered to replace an RNA stem loop known to regulate tombusvirus replication leading to theophylline dependent viral titers (Wang and White 2007). Conformational switching riboswitches capable of controlling gene expression *in vivo* have also been engineered to utilize the theophylline aptamer (Suess, Fink et al. 2004; Kim, Gusti et al. 2005). Riboswitches are different from the antiswitch because they work *in cis*, controlling expression from the mRNA sequence of the gene (Reviewed in Edwards, Klein et al.

2007). Self cleaving (Tang and Breaker 1997; Soukup and Breaker 1999) or ligating (Robertson and Ellington 2000) ribozymes have been rationally engineered and experimentally selected to incorporate the theophylline aptamer for allosteric regulation.

In addition fluorescent aptamer beacons have been designed around the theophylline aptamer. The theophylline dependent ribozyme, created by Soukup and colleagues, was engineered to separate a fluorophore and quencher (Frauendorf and Jaschke 2001) upon cleavage. Likewise, Stojanovic and colleagues created modular sensors using both a malachite green and theophylline aptamer (Figure 1.4C; 2004), in which dye binding was dependent on conformational changes induced through the binding of theophylline. Similar to work presented here, a theophylline biosensor that activates a molecular beacon has previously been developed rationally (Rankin, Fuller et al. 2006), however the utility of our method is not to detect thrombin, but rather a study of the energetics involved in and the generation of generic antiswitches.

Whereas it is generally difficult to model aptamer ligand interactions *a priori*, due to noncanonical base pairs and unusual structural motifs often found in aptamer ligand interactions (Feigon, Dieckmann et al. 1996; Patel 1997), through computational RNA structure prediction, we were able to search vast sequence libraries for known structural motifs. Using ViennaRNA (Hofacker 2003) and Perl scripts we have computationally created and sieved a pool of random hairpins with unique sequence composition and length that, when appended to the theophylline aptamer, were capable of activating molecular beacons experimentally in the presence of theophylline. We also show that tuning either the beacon or aptamer thermodynamic parameters provides higher signal to noise ratios and increased kinetic rates of activation. It appears that the binding of the antiswitch to theophylline and consequently the molecular beacon irreversibly stabilizes the antiswitch into an “active” conformation even upon the dissociation of theophylline.

RESULTS AND DISCUSSION

Development and computational assessment of the antiswitch model

In an attempt to create a modular system for the development of signaling aptamers, we chose to model an antiswitch originally designed and published Bayer and colleagues whereby the binding of a cognate analyte (theophylline) induces a switch stabilizing the 3' stem of the aptamer and denaturing the 5' stem. This single stranded 5' stem is then free to bind a target molecular beacon loop sequence **Figure 4.3**. The antiswitch model tacitly assumes there are six possible conformational states the antiswitch and beacon could be within depending on the energetics and presence of theophylline or molecular beacon (**Figure 4.3A-F**). These states are controlled by equilibrium constants for the conversion between states, and is specified by the relative energetic stabilities of each state and the binding of aptamer with analyte.

The antiswitch starts in a stable equilibrium with a molecular beacon termed the inactive unbound "OFF" state since it cannot actively hybridize the molecular beacon or bind the target (Figure 4.3 A). Through conformational switches induced in part by target and molecular beacon, the antiswtich can become active (capable of binding the molecular beacon; **Figure 4.3B, C, E, F**), bind the ligand, (bound, **Figure 4.3D, E, F**) and hybridize to the beacon, which subsequently increases the fluorescent signal ("ON", **Figure 4.3C, F**). It is necessary to design an antiswitch, in which the inactive conformer is stable when a molecular beacon is present so hybridization doesn't significantly drive the equilibrium to the active "ON" conformation independent of ligand binding. An antiswitch that posses an overly stable inactive conformer however, may not be capable of conformational switching upon ligand addition or assume sufficient alternate

conformations at equilibrium. If however, the inactive unbound conformer is unstable and the active aptamer is most prevalent at equilibrium, then there will be high background. In addition, if target binding does not drive the equilibrium towards an active state where the 5' hairpin is open and available, then target depending signaling will not be detected. Ultimately, there must exist a coupled equilibrium between the antiswitch and the beacon that must be taken into account.

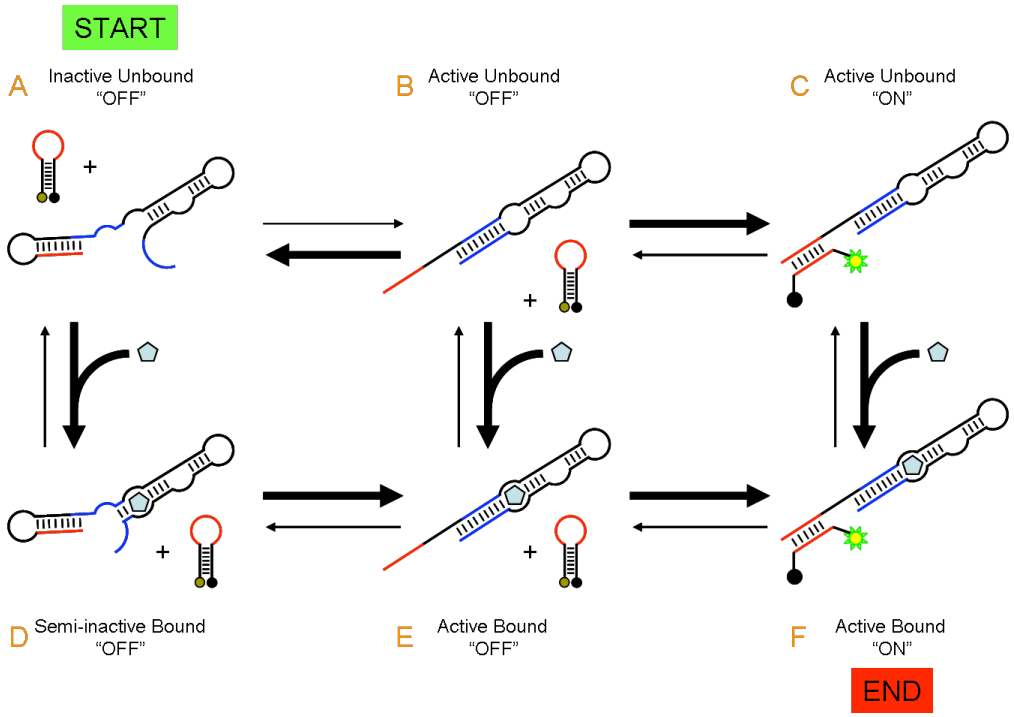


Figure 4.3. Kinetic pathways of Antiswitch activation by target binding and molecular beacon signaling. The states are numbered in orange. Starting in an inactive conformer (A) , the antiswitch should remain at equilibrium separate from a molecular beacon in the absence of target. Upon ligand addition, a conformational change occurs in the antiswitch to the active conformer capable of binding both the target and the molecular beacon (F).

To begin selecting new antiswitch constructs, we analyzed the energetics of the previously published antiswitch components to determine the parameters that may prove useful in switching. Ultimately, the previous antiswitches maintained an inactive yet

stable structure *in vivo* and were capable of conformation switching into the active structure stabilized not only by the addition of theophylline (binding free energy = -8.96 kcal/mol; Gouda, Kuntz et al. 2003) but also by hybridization to an area surrounding the AUG start codon of the mRNA expressing green fluorescent protein. In this manner, GFP expression was dependent on theophylline availability and constitutively expressed in yeast (**Figure 4.4**).

S1 was the original construct developed by Bayer *et al* which displayed switching *in vivo* at theophylline concentrations around 1mM. When the 5' antiswitch hairpin was destabilized (S4 construct), switching occurred at 100 fold lower theophylline concentrations. However when the hairpin was stabilized, theophylline dependent switching didn't change significantly from that of the original construct S1. This suggests that S1 is actually quite stable within the *in vivo* system and destabilizing it may provide lower detection limits, increased speed or a favorable changes in background signal. S1 was predicted to have a MFE with a ΔG of (-24 kcal/mol), yet destabilizing the 5' loop decreased the overall free energy to -20.5 kcal/mol. The core portion of the aptamer present in both the active and inactive structure has a minimum free energy of -6.2 kcal/mol (**Figure 4.4 purple box**), whereas the 3' antiswitch tail hybridizes with a portion of the 5' hairpin in the active confomer possessing a minimum free energy of -10.8 kcal/mol (**Figure 4.4 blue sequence**).

The authors noted that the energy of the 3' tail when bound in the active conformer (-10.8 kcal/mol) combined with the binding energy provided by theophylline (-8.9 kcal/mol) was identical to the folding energy of the 5' hairpin (-17.8 kcal/mol). In addition, we noticed that the destabilized antiswitch MFE of S4 (-20.5 kcal/mol) was nearly identical to the MFE of the aptamer structure in the active bound conformer (-20.2 kcal/mol). Bayer and colleagues also designed tetracycline dependent antiswitches. We

compared the structural and energetic profiles within this group of antiswitches to the published *in vivo* expression levels and found similar correlations to the theophylline dependent antiswitches (data not shown).

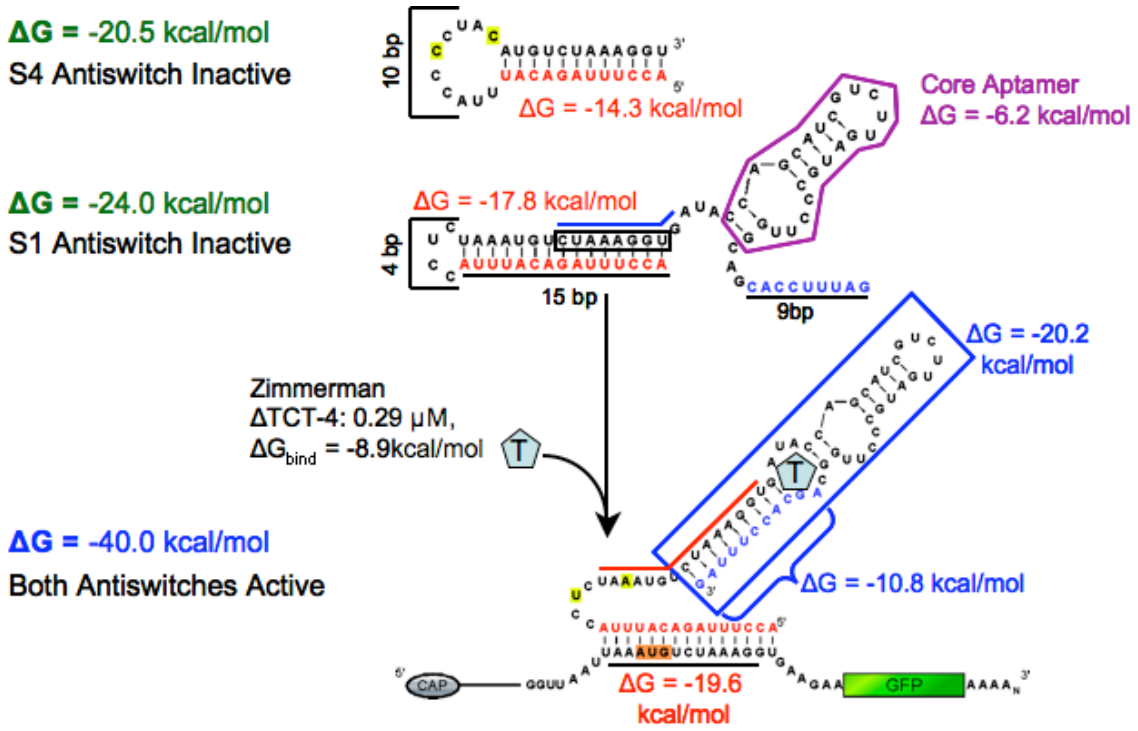


Figure 4.4. S1 and S4 antiswitch activation by theophylline (pentagon T). The MFE antiswitch is labeled in green. The MFE for the 5' hairpin is denoted with red. The core aptamer is boxed in purple. The core aptamer in the active bound conformer is boxed in blue. The red and blue lines denote, in which the hairpins hybridize in the alternate conformer. The active antiswitch binds to the AUG start codon of the GFP gene transformed into yeast cell line. The yellow highlighted bases represent differences between S1 and S4.

Computational selection of antiswitches

After assessing potential parameters from the published antiswitches, a computational selection was performed (see Materials and Methods) to determine potential antiswitches from a pool of random sequences. We first developed a program (see Computational selection in Materials and Methods) to create a pool of random

sequence hairpins (**Figure 4.5**). The pool size, hairpin length and loop size were all input variables. The use of a 10,000 sequence pool often provided ample variation to selected representative variants within relatively short time frames (10 minutes to 3 hours depending on stem length). Each random sequence was added to the pool array if it contained no internal structure or sequence runs (listed in Materials and Methods) until the pool reached the predefined size. For the initial pool, a hairpin length of 15 nucleotides was utilized. If the sequence did not pass the tests, it was deleted and a new sequence was created. Once the pool was complete, each member was converted into a hairpin by adding a loop sequence of CCUC and the reverse complement of the pool sequence. The hairpin was also tested to verify it indeed formed the proper stem loop structure.

Following this, the actual antiswitch was constructed by appending the hairpin to the 5' portion of the aptamer sequence along with a 10 nucleotide 3' tail of complement sequence and predefined length capable of hybridizing a portion of the hairpin. We also searched the complete antiswitch sequence for repeats. In addition the antiswitch was folded to determine if the proper MFE structure formed. The MFE structure was defined as one, in which the 5' hairpin (**Figure 4.5, red box**) and the core aptamer (**Figure 4.4, purple box**) formed, and the 3' tail remained unpaired (**Figure 4.5, blue sequence**).

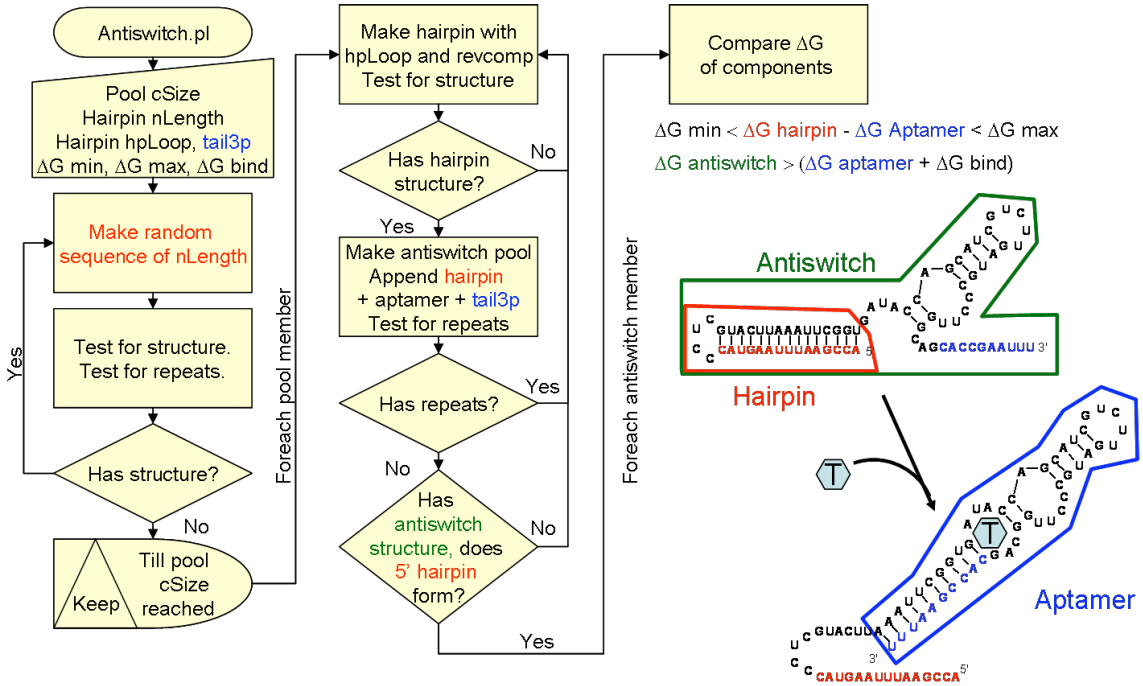


Figure 4.5. Flowchart of computational selection. The MFE for the antiswitch present as the inactive unbound "OFF" conformer (green box) switches into a the active bound "ON" conformer (blue box) upon the addition of target (T hexagon) and molecular beacon. The 5' hairpin is boxed in red.

Lastly free energy profiles of each pool member were calculated. Those sequence variants, in which the minimum free energy difference between the hairpin (red box) and aptamer (blue box) were within a range of 2 and 3 kcal/mol, were kept. Those members that fell outside this range were discarded (**Figure 4.5**). In addition, we verified that the folding free energy of the antiswitch was less stable than the combination of the folding free energy of the active unbound "OFF" conformer and the binding energy provided by the ligand.

Since S1 had an antiswitch stem length of 15 base pairs and a loop sequence consisting of "CCUC", we started with these variables. Also, we were uncertain whether a large or small difference between the inactive antiswitch structure and the active antiswitch structure was important so we chose a sequence representing a larger

difference (similar to S1) to be certain that the inactive conformer would predominate at equilibrium.

Total Seq	Name	BP	loop	hp	3'p tail	Antiswitch MFE	Aptamer MFE	Hairpin MFE	Antiswitch -Apt	Apt -HP	Activity
Bayer	S1	74	4	15	9	-24.00	-20.20	-17.80	**-3.80**	-2.4	not active
Bayer	S4	74	10	12	9	-20.50	-20.20	-14.30	**-0.30**	-5.9	2.1 fold
randpool	AS1	75	4	15	10	-25.10	-21.00	-18.90	**-4.10**	-2.1	not active
1083	A6-10	63	4	10	11	-21.90	-19.00	-15.70	**-2.90**	-3.3	Slow, 3 fold
1083	A6-11	63	4	10	11	-23.10	-23.10	-16.90	0.00	-6.2	Fast, 1.5 fold
3454	A8-6	70	6	12	12	-22.80	-22.80	-16.60	0.00	-6.2	Fast, 2 fold

Table 4.1. Energetics of the published antiswitches (S1 and S4) along with new designs (AS1, A6-10, A6-11, A8-6). The colors are consistent with Figure 4.3 and Figure 4.4.

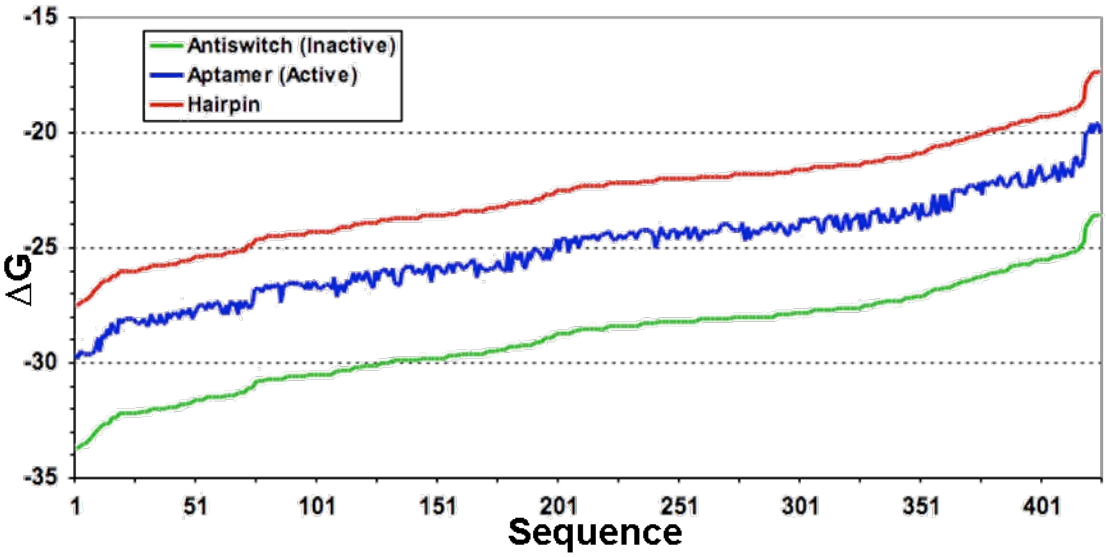


Figure 4.6. Energetic profile for a pool of antiswitches. The sequences that conform to the parameters defined in (Materials and Methods) similar to S1 in a 10,000 variant pool consisting of a 15-mer hairpin length with a 4 base pair loop and 9 bp 3' tail. The cutoff difference between the antiswitch and aptamer was 3 kcal/mol. The previously selected antiswitch S1 is labeled with an arrow.

From an original random pool of 10,000 variants that was winnowed down to 425 members that conformed to the parameters, we selected one variant (AS1; **Table 4.1**) that

most closely matched the energetic profile of S1. In addition two of the constructs published by Bayer and colleagues (S1 and S4; **Table 4.1**) were synthesized and tested for preliminary activation *in vitro*. Since S1 showed relatively higher free energies compared to the other members of our *in silico* pool, AS1 was selected in this energetic area of the pool (**Figure 4.6**). It should be noted that, the difference between the MFE of the active conformer and the minimum free energy of the inactive conformer is larger with S1 and AS1 compared to S4 (**Table 4.1**).

We also designed two molecular beacons to detect theophylline dependent conformational switching of the antiswitch. The first was similar a portion of the GFP 5'UTR previously utilized (Bayer and Smolke 2005) and could hybridize with both S1 and S4. The second was specific for AS1.

AS1 theophylline dependent activation

We synthesized and purified the variant (AS1) selected computationally along with Bayer's antiswitches and the molecular beacons. We then assayed fluorescent activation of the respective molecular beacon with each of the constructs (**Figure 4.7A**). Ligand-dependent fluorescence activation was measured by adding 250uM theophylline (100 fold higher than the published aptamer K_d) and the results were compared with activation data from S1 and the destabilizing mutant S4 (**Figure 4.7A**). Unfortunatly, we could not detect theophylline dependent activation from either S1 or AS1, however a 2.3 fold signal increase was detected with S4. This suggests that the inactive conformers of both AS1 and S1 are overly stable and cannot be switched into active conformers upon ligand addition. In addition, S4 does not fully activate the beacon. It could either be that the aptamer cannot bind due to the stability of the inactive conformer, or that binding does not induce a sturctural change capable of displacing the 5' hairpin.

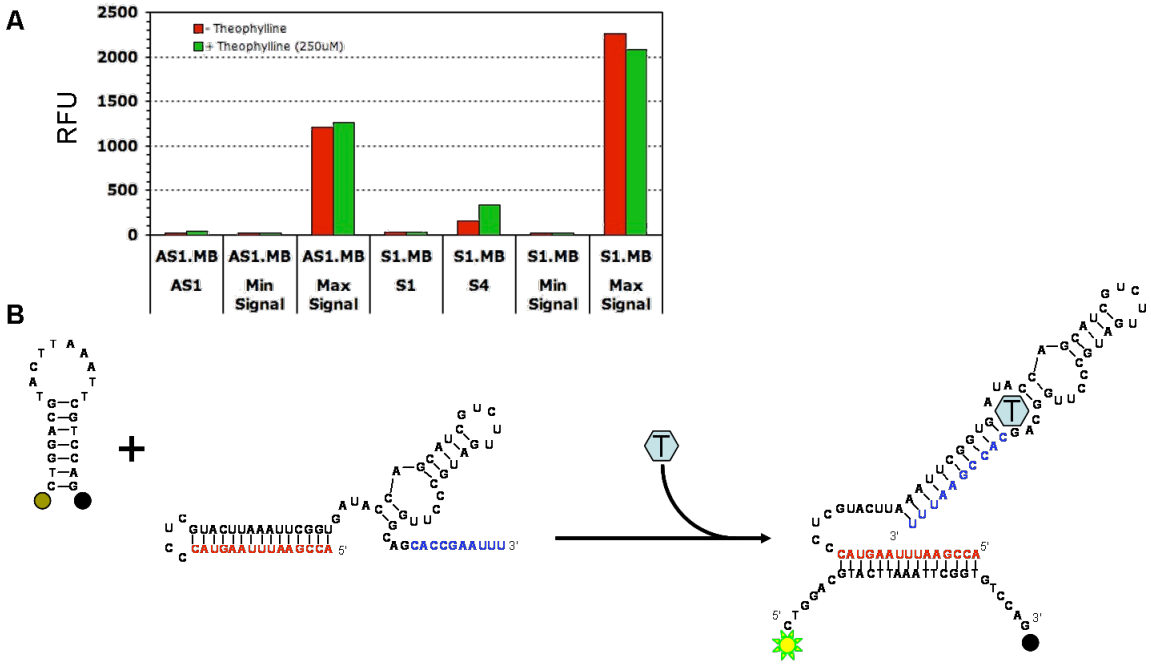


Figure 4.7. Molecular beacon based activation of assay of AS1, S1 and S4. A.) The fluorescent activation of AS1 (test) compared with a positive control, maximum and minimum signal is shown. These are performed in conjunction with tests of S1 and S4 including the minimum and maximum beacon fluorescent. Concentrations of 600nM antiswitch, 300nM molecular beacon, 600nM activators, 250uM theophylline 20 minutes at 37°C. B.) Structural switching of AS1 with the beacon activator and the antiswitch AS1.

Computational selection of A6-10, A6-11 and A8-6

Whereas the previous set of antiswitches were validated by *in vivo* expression, we found the same constructs were unable to hybridize with a molecular beacon *in vitro* thus limiting their usefulness as reagentless diagnostics. In addition, the parameters utilized for modeling AS1 around the energetics of S1 provided an antiswitch trapped in the inactive unbound conformer. Therefore, we thermodynamically modeled three new antiswitch constructs with less stable 5' hairpin energetics (**Figure 4.8A and Table 4.1**).

We again tested these constructs in the presence of theophylline except we increased the concentration of theophylline, and utilized equimolar concentrations of

beacon and antiswitch instead of a two-fold difference before. The beacon and antiswitch were incubated in the absence of theophylline until a stable equilibrium was achieved, then 1mM theophylline was added and the reaction was monitored at various time points over one hour (**Figure 4.8B**).

It is apparent that the overall activation level is partially determined by the stability of the molecular beacon alone. Both A6-10MB and A6-11MB have a predicted MFE of around -5.9 kcal/mol, whereas A8-6MB is much less stable at only -1.9 kcal/mol. When comparing the percent total signal of the beacons in the absence of antiswitch, A8-6 has a predictably higher signal than either of the other two beacons. The percent total signal is calculated by dividing the signal at each time point by the maximum possible signal measured from a fully denatured beacon.

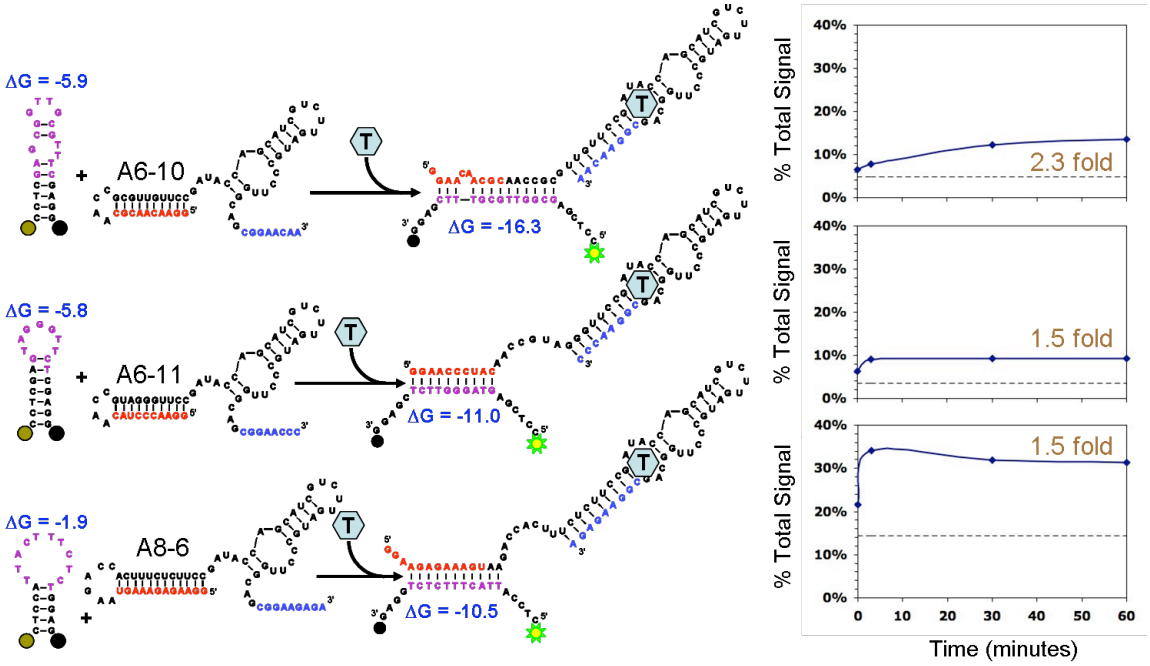


Figure 4.8. Prediction of functional antiswitches A6-10, A6-11 and A8-6. A.) A structural model for antiswitch activation of three constructs with different 5' hairpin sequences. Similar color conventions are utilized. B.) Activation assay showing activation over time. Equimolar concentration of antiswitch and molecular bacon (500nM) were incubated with 1mM Theophylline (red, green yellow bars). A6-10 slow, 2.1 fold activation. A6-11 Fast but 1.5 fold activation. A8-6 Fast, 1.6 fold activation, beacon less stable.

There also appears to be a correlation between the stability of the beacon bound to the antiswitch and the kinetic rate. A6-11MB and A8-6MB have similar beacon hybridization free energies (-11.0 kcal/mol and -10.5 kcal/mol respectively) and fully activate within the three minutes between theophylline addition and the first fluorescence measurement as opposed to A6-10 and S4 (-16.9 kcal/mol and -17.4 kcal/mol), in which the signal was still increasing 60 minutes after theophylline incubation.

Lastly, the data suggests that the stability of the 5' hairpin in the inactive unbound antiswitch conformer affects the kinetic rate as suggested by previous results comparing S1 with S4. This apparent difference in kinetic rates could be due to the relative stabilities within the antiswitch of A6-11 and A8-6 (**Table 4.1**) since there is no predicted

minimum free energy difference between the inactive unbound antiswitch and the aptamer portion of the active conformer.

Specificity of A6-10 and A8-6

We next wanted to determine how specific the antiswitch was for the target theophylline. It has previously been shown that the unmodified aptamer can discriminate between theophylline and caffeine by up to 10,000 fold, yet the aptamer is capable of binding caffeine at relatively high ligand concentrations above 5mM (**Figure 4.2**; Zimmermann, Wick et al. 2000). By incorporating the aptamer into an antiswitch, we wanted to verify that these changes didn't interfere with ligand specificity. We therefore tested the antiswitch constructs A6-10 and A8-6 for activation in the presence of ligand concentrations from 10nM to 1000nM (**Figure 4.9**). As can be seen, there was relatively little additional fluorescence change in the presence of caffeine compared to no ligand. Surprising, however theophylline could be detected as low as 10uM even though this is much lower than the available antiswitch (500uM A6-10 and 250uM A8-6). These results suggest that once inactive unbound antiswitch switches into an active conformer upon ligand binding, the active conformer becomes energetically trapped by hybridizing to the beacon ("ON"), although theophylline subsequently dissociates.

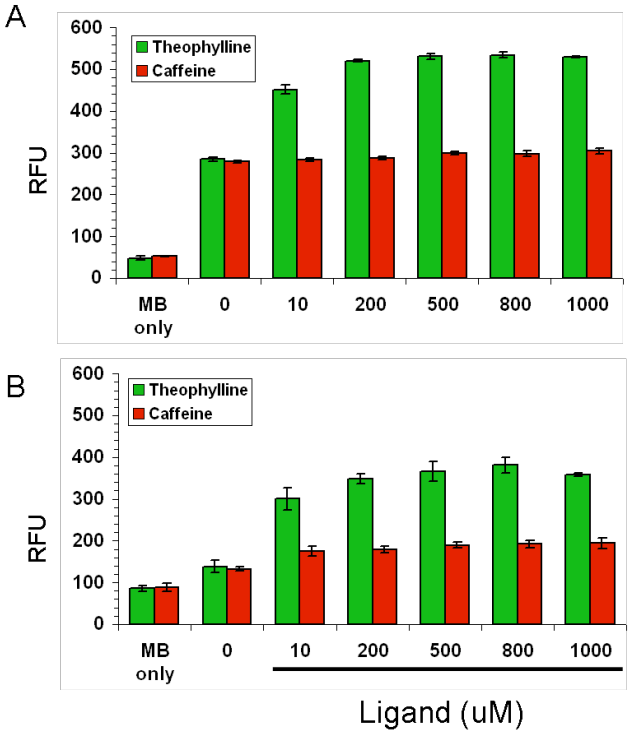


Figure 4.9. Specificity of A6-10 and A8-6. A.) A6-10 tested (500nM) at various concentrations of theophylline or caffeine and measured after 1.5 hours incubation. B.) A8-6 (200nM) assayed identically to A6-10.

Improving kinetic response by changing the beacon energetics

We wanted to determine whether we could increase the speed of activation through manipulation of the beacon alone. Whereas A6-10 and the beacon A6-10MB displayed the highest fold activation, it proved the slowest to activate. We realized that upon folding, the loop of A6-10MB was closed down by two stacking G:C base pairs. We therefore, designed two new beacons specific for A6-10 with larger hairpin loops, (**Figure 4.10A**).

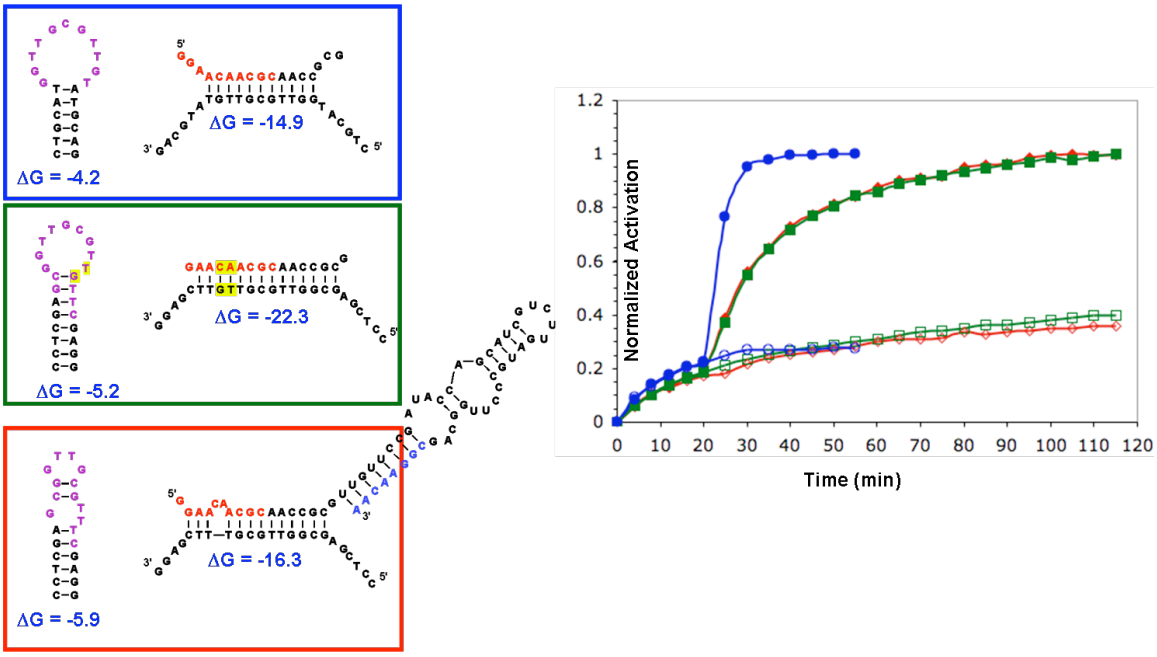


Figure 4.10. Improving the kinetics of antiswitch activation. A.) Two new beacons (green box and blue box) were assayed along with the previous beacon (red box) all specific to A6-10. B.) Kinetic assay for the modified beacons. Result colors matched the molecular beacon in (A). The antiswitch and respective beacons were mixed at equimolar concentration (500nM) and assayed in the absence (open markers) or presence of 500uM theophylline (closed markers). Activation was normalized to the minimum signal (at time point 0) and maximum detectable signal upon ligand addition.

The antiswitch was assayed with the three beacons separately over time in the presence of 500uM theophylline or caffeine (**Figure 4.10B**). Again, in all cases, there was no difference between the addition of caffeine and the absence of target suggesting the antiswitch was specific for theophylline (data not shown). The new beacon (A6-10.8TMB) designed with the largest loop and least stable free energy (ΔG = -4.3 kcal/mol) also provided the fastest activation signaling to maximum levels within 10 minutes, whereas the other two beacons showed nearly identical kinetic rates taking nearly 2 hours to fully activate.

CONCLUSIONS AND PROSPECTS

We recently began work on this project and a number of experiments are still necessary. For instance, we would like to test the limits of detection for theophylline. A substantial increase in fluorescence was observed from the addition of antiswitch alone (**Figure 4.8B**) and thus predicting antiswitches with more stable hairpins may be necessary to reduced background signal.

Once the proper energetic parameters have been fully elucidated, we would like to detect more complex targets by incorporating aptamers capable of binding proteins or whole cells. However based on results with anti-ricin and anti-yop51 aptamers, choosing an aptamer, in which the structure has been experimentally determined may be an important factor. Lastly, much of the data must be reacquired to calculate statistical deviation.

MATERIALS AND METHODS

Synthesis and Purification

Some of the initial constructs were ordered through IDT (Coralville, IA), whereas we synthesized other including all of the beacon activators and some of the molecular beacons. We found no difference in experimental data between the two sources (data not shown) and only time or cost determined the source. Materials were synthesized as in (Hall, Hesselberth et al. 2007). Briefly 5' 6-FAM moieties were introduced by 5' phosphoramidite chemistry, whereas 3' quenchers were incorporated through controlled pore glass (CPG) solid supports (Glen research, Sterling, VA) during synthesis. The quenchers utilized include dabcyl (4-(4-dimethylaminophenylazo)benzoic acid) and Black Hole Quencher 1 (BHQ-1). After deprotection samples were lyophilized and

purified by HPLC (Nutiu and Li 2003) with the following changes: column was an XTerra MS C18 Column with dimensions of 4.6mm x 30mm and 2.5um bead diameter. The two solvents were 0.1M TEAA and 100% acetonitrile on a nonlinear gradient from 5% to 35% acetonitrile over 30 minutes at a flow rate of 3 mL/min. We collected 2/3 of the main peak width showing absorption at both 260nm and 495nm. The purified DNA was lyophilized and resuspended in 100ul diH_2O.

RNA aptamers were transcribed using AmpliScribe T7 high yield transcription kits (Epicentre, Madison WI) from 200pmol dsDNA by first extending a pair of overlapping oligonucleotides with SuperScript II (Invitrogen, Carlsbad CA). The oligonucleotide sequences utilized for RNA production are listed in **Table 4.2**. RNA was purified on an 8% denaturing polyacrylamide gel. For all sequences, strand concentrations were detected by UV spectroscopy at 260nm on a Nanodrop ND-1000 (Wilmington, DE) using extinction coefficients and adjusted to 10uM for each oligonucleotide sequence.

Kinetic Assays

Assays were performed with equimolar concentrations of beacon and RNA antiswitch (500nM) in 50ul reactions at room temperature unless otherwise specified. Briefly, the aniswitch and beacon were separately mixed in signaling buffer (100mM HEPES pH 7.3, 50mM NaCl, 5mM $MgCl_2$ (Jenison, Gill et al. 1994)), heat denatured at 60° C for 3 minutes, cooled to 25° C at 0.2° C/sec. They were allowed to further thermally equilibrate at room temperature (26° C) for 5 to 15 minutes then mixed. Assays were carried out in Corning 96-well ½ area black opaque plates (Corning, NY). Maximum fluorescence was measured by hybridizing the beacon with an unmodified complement sequence, whereas minimum fluorescence was measured with the beacon alone. Aptamers were assayed for activation in the presence or absence of 1mM

theophylline or caffeine (Sigma, St. Louis, MO) on a BioTek Synergy HT (Winooski, VT) fluorescent plate reader at between 25.7° C and 26.8° C with a gain of either 48 or 50 such that a buffer blank yielded a signal of 26 RFU.

Computational methods

The free energies of RNA secondary structures were calculated using the ViennaRNA package (Hofacker, Fontana et al. 1994). Scripts to automate the evaluation and comparison of RNA secondary structure profiles were coded in the scripting language Perl (http://www.perl.org/). These programs first generated nucleic acid antiswitch sequences with randomized 5' hairpins and 3' tails. For each sequence in the randomized population, the free energy of folding was calculated and the sequence was also compared to a list of sequence repeats (AAAA, GGGG, UUUU, CCCC, AUAUA, UAUAU, GCGCG, CGCGC, GUGUG, UGUGU, UUGGU, GGUUGAUAUU, UAUAA, AUUAA, UAAUU, AAUUA, UUAAU). RNAfold was utilized with the following constraints: "noLonelyPairs" that removes one base pair loop energies from the calculation of minimum free energy, "d = 0" that similarly prevents dangle end energies, "tetra_loop" to incorporate known stable tetra loops, and "no_closingGU" such that a GU pair cannot close a loop. All of these constraints were suggested in the ViennaRNA manuals. Refer to **Figure 4.5** for a simplified flowchart of the program.

Name	Sequence
A6-10f	GATAATACGACTCACTATAGGAACAACGCAACCGCGTTGTTCCGATACCAGCATCG
A6-10r	TTGTTCCGCTGCCAAGGGCATCAAGACGATGCTGGTATCGGAACAACGCGGTT
A6-10MB	/56-FAM/CCTCGAGCGGTTGCGTTTCGAGG/3Dab/
A6-10MBrc	CCTCGAAACGCAACCGC
A6-10.2MB	/56-FAM/CCTCGAGCGGTTGCGTTGTTCGAGG/3Dab/
A6-10.8MB	/56-Fam/CTGCATGGTTGCGTTGTTATGCAG/3Dab/
A6-10.8TMB	/56-Fam/CTGCATGGTTGCGTTGTATGCAG/3Dab/
A6-11f	GATAATACGACTCACTATAGGAACCCTACAACCGTAGGGTTCCGATACCAGCATCG
A6-11r	GGGTTCCGCTGCCAAGGGCATCAAGACGATGCTGGTATCGGAACCCTACGGTTG
A6-11MB	/56-FAM/CCTCGAGTAGGGTTCTCGAGG/3Dab/
A8-6f	GATAATACGACTCACTATAGGAAGAGAAAGTAAGACCACTTTCTCTTCCGATACCAGC
A8-6r	TCTCTTCCGCTGCCAAGGGCATCAAGACGATGCTGGTATCGGAAGAGAAAGTGGTC
A8-6MB	/56-FAM/CTCCATTACTTTCTCTGGAG/3Dab/
S1f	TTCTAATACGACTCACTATAACCTTTAGACATTTACCTCTAAATGTCTAAAGGTGATAC
S1r	CTAAAGGTGCTGCCAAGGGCATCAAGACGATGCTGGTATCACCTTTAGACATTTAGAGG
S4f	TTCTAATACGACTCACTATAACCTTTAGACATTTACCCCTACATGTCTAAAGGTGATAC
S4r	CTAAAGGTGCTGCCAAGGGCATCAAGACGATGCTGGTATCACCTTTAGACATGTAGGGG
S1.mb	CTGGACGTAAATGTCTAAAGGTGTCCAG
S1.mb.act	CTGGACACCTTTAGACATTTAC
AS1f	TTCTAATACGACTCACTATA
AS1r	AAATTCGGTGCTGCCAAGGGCATCAAGACGATGCTGGTATCACCGAATTTAAGTACGAGGGTACTTAAAT
AS1.mb	/56-FAM/CTGGACGTACTTAAATTCGGTGTCCAG/3BHQ_1/
AS1.mb.act	ACCGAATTTAAGTAC
AS1.refold	AATTTAAGCCACTAT

Table 4.2. Sequences utilized in this chapter.

Chapter 5: Kinetic optimization of a protein-responsive aptamer beacon

INTRODUCTION

Aptamers have proven to be useful as biosensors, in large measure because they can be readily adapted to a variety of sensor platforms and signal transduction schemes through the rational manipulation of their sequences and secondary structures (reviewed in Nutiu and Li 2005; Bunka and Stockley 2006; Deisingh 2006; Cho, Rajendran et al. In Press). One highly generalizable scheme for the generation of so-called aptamer beacons involves generating antisense oligonucleotides that denature the functional structure of an aptamer, and then placing fluorescent reporters and quenchers on the sense and antisense strands (Nutiu and Li 2003). Upon binding to a target ligand, the equilibrium between the antisense-bound, non-functional structure and the free functional structure is altered, with a concomitant change in observed fluorescence. However, while this and other signal transduction schemes readily yield aptamer biosensors that can sensitively detect a variety of analytes, the kinetic properties of the resultant biosensors have not been examined in any detail.

We have therefore generated a series of aptamer beacons that can detect the protein thrombin, and examined their kinetics of signaling. While the anti-thrombin aptamer is generally chosen for biosensor development because of its simplicity, in this instance it was chosen in large measure because so many previous studies have used this aptamer. The original thrombin aptamer was a ssDNA with a consensus sequence of GGtTGG(N2-5)GGtTGG (Bock, Griffin et al. 1992). It was not only the first time a ssDNA selection had successfully been completed, but the first time a target without native nucleic acid binding properties was shown to bind DNA. In 1993, the structure of

the best aptamer d(GGTTGGTGTGGTTGG) was elucidated by two dimentional NMR (Macaya, Schultze et al. 1993) three dimentional NMR (Wang, McCurdy et al. 1993; Schultze, Macaya et al. 1994) and X-Ray crystallography (Padmanabhan, Padmanabhan et al. 1993). These studies showed that the potassium dependent aptamer formed a unimolecular quadruplex in solution where two G-quartet stacks were connected by two T-T loops and a TGT loop (**Figure 5.1A, black underlined**). They also found that the aptamer binding site overlaps with thrombin's fibrinogen recognition exosite (FRE) along with the thrombin platelet receptor and thrombomodulin binding sites.

Following this, both RNA (Kubik, Stephens et al. 1994) and additional DNA selections (Macaya, Waldron et al. 1995; Tsiang, Jain et al. 1995; Tasset, Kubik et al. 1997) were completed. While each of the selected aptamers contained a G-quadruplex, the subsequent DNA selections suggested an additional four to seven base pair duplex could help stabilize the quadruplex and lower the binding affinity (**Figure 5.1C vs 1A**). It was also proposed by Macaya *et al* that this duplex could provide additional contacts with thrombin. While the RNA aptamer was shown to bind thrombin's heparin-binding exosite, all DNA aptamers other than those selected by Tasset *et al* bound the FRE. Tasset and colleagues selected a ssDNA aptamer termed 60-18 with the sequence 5'-AGTCCGTGG**TA**GGGCAGGTTGGGGTGACT-3' that displayed potassium independent binding similar to other aptamers with extended duplex (**Figure 5.1C**). In contrast however, the bolded "TA" in the previous sequence was found through photo-crosslinking experiments to direct the aptamer to the heparin-binding exosite while a TT loop directs the aptamer to the FRE site in previously selected DNA aptamers.

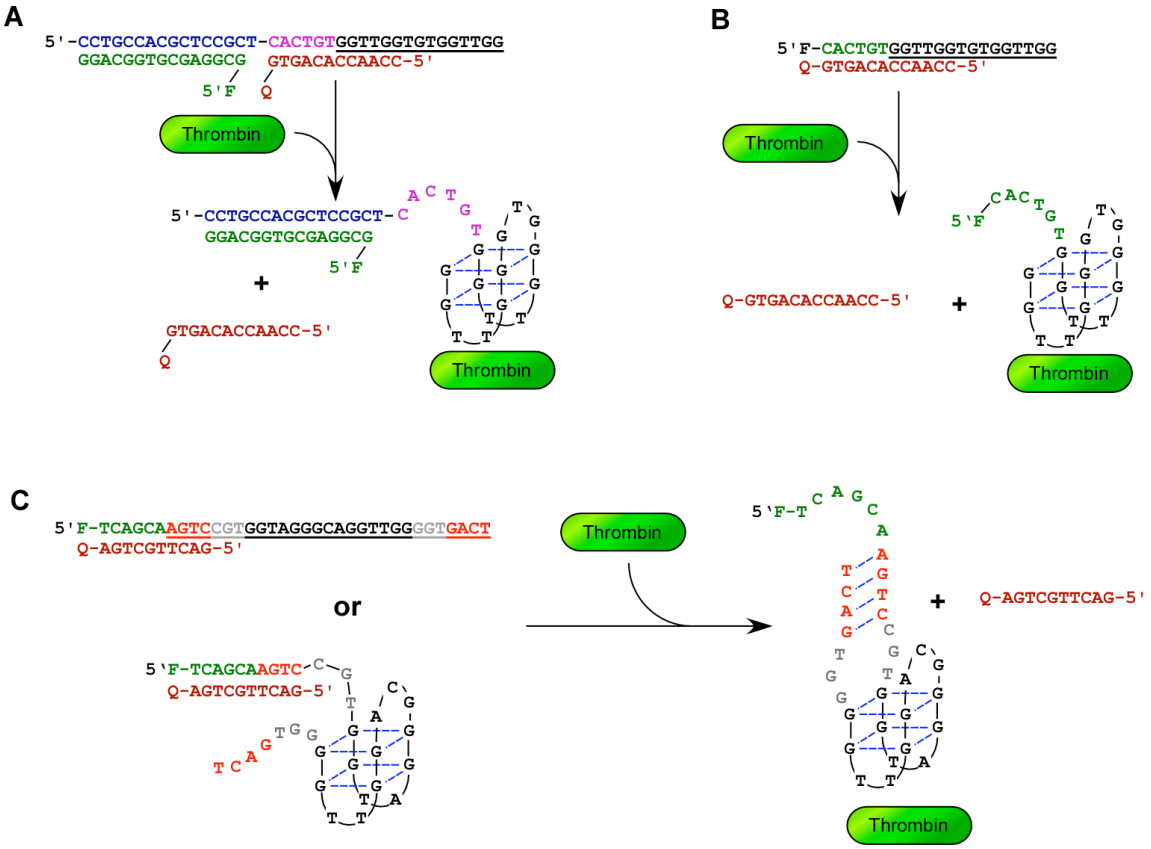


Figure 5.1. Structure Switching Models. A.) Structure-switching reporter developed by Nutiu (2003). Blue is Stem-1, purple is Stem-2, underlined sequence is the original DNA sequence consensus for the thrombin binding aptamer (Bock, Griffin et al. 1992), later called G15D (Tasset, Kubik et al. 1997). Green is FDNA1 with a 5'-Fluorescein and dark red is QDNA6 with a 3'-DABCYL. The G-quartet is drawn to represent three dimensional structural arrangement. B.) Representative activation of Apt4 (Stem-2) and Th.Q1 (QDNA6). C.) AptH complexed with Th.Q-5-2. AptH is derived from 60-18[29] (underlined; Tasset, Kubik et al. 1997) with TCAGCA appended to the 5' end. Similar sequence and color conventions are used where the green sequence is bound to a fluorescein, the red sequence represents a duplex stem in the aptamer conformer, grey is an internal bulge characteristic of this aptamer.

Once the structure and functional characteristics of the thrombin aptamer were elucidated, rationally designed signaling biosensors were developed. In 2001, Hamaguchi *et al* re-engineered the original thrombin aptamer selected by Bock *et al* to act as a signaling aptamer by including additional bases at the terminal 5' end

complementary to those at the 3' end (Hamaguchi, Ellington et al. 2001). This sequence had a minimum free energy structure that formed a hairpin-loop akin to a molecular beacon in the absence of thrombin. When thrombin was added, the G-quadruplex forced a global conformational rearrangement separating the fluoresecent moieties affixed to the 5' and 3' ends. It is interesting to note that the biosensor could switch structures quickly depending on the presence of divalent magnesium (beacon) or monovalent potassium (G-quartet). It was also shown that the thrombin aptamer could act as a FRET beacon without the additional sequence used by Hamaguchi *et al* (Li, Fang et al. 2002). In the absence of thrombin the aptamer was generally unstructured thus separating fluorescein and DABCYL appended to the 5' and 3' ends, respectively. Thrombin dependent structural stability allowed the terminal fluorophores to interact thus changing the fluorescent spectra.

In an attempt to generalize the design of aptamer biosensors, Nutiu and Li developed both a tripartite and bipartite signaling aptamer (Nutiu and Li 2003). They called this a "duplex-to-complex structure-switching approach," whereby the global conformational changes characteristic of thrombin binding led to the displacement of a hybridized quencher oligo. The bipartite method was developed with the ATP aptamer 5'-ACCTGGGGGAGTATTGCGGAGGAAGGT-3' (Huizenga and Szostak 1995) was extended at the 5'end and a fluorephore was added, a secondary oligo with a quencher at the 3'end was designed to hybridize to the aptamer extension and disrupt the native aptamer structure in the absence of target (**Figure 5.1A**). Once ATP was added, the aptamer would undergo a conformational change that would displace the quencher oligo. The tripartite method was used to design biosensors around the ATP and thrombin aptamers. It is very similar to the bipartite model yet more modular. Instead of modifying the aptamer with a fluorophore and using a quencher oligo, the aptamer has an

additional extension that is complementary an oligo with a fluorophore at it's 5' end. The complementary oligos are separated such that the labels are in close proximity. Upon ligand binding, the quencher oligo is displaced, yet the fluorescinated oligo is left hybridized, thus yielding an increase in fluorescent signal. Activation was measured both as a function of temperature and in varying magnesium concentrations. The aptamer activated nearly instantaneously at 37C, yet was considerably slower at lower temperatures. In addition to temperature variations, others have characterized binding kinetics as a function of pH, ionic strength and aptamer configuration using quartz crystal microbalance for detection.(Hianik, Ostatna et al. 2007). Both aptamers were used in a proximity assay that allowed coassociation of the aptamers to bring together two "signaling oligonucleotides" thus producing a large change in FRET signal (Heyduk and Heyduk 2005).

Herein, we developed two different bipartite thrombin signaling aptamers with fast activation rates at room temperature (25° C) and provided guidelines for the design of quencher oligos and requisite quencher moieties. By examining the relationship between the design of biosensors based on the anti-thrombin aptamer and their response dynamics, we hoped to provide a foundation for the comparison and understanding of many different sensor platforms and methods.

RESULTS AND DISCUSSION

Designing aptamer beacons that signal the presence of thrombin

Nutiu and Li (2003) have previously generated antisense displacement aptamer beacons for the detection of thrombin. In this work, two antisense oligonucleotides containing fluorescent and quencher moieties were annealed to and thereby denatured a

short, anti-thrombin DNA aptamer known to form a quadruplex structure (Clone 29; Bock, Griffin et al. 1992). Interaction with thrombin disfavors the tripartite, quenched complex and stabilizes the quadruplex structure, thus destabilizing the quencher oligonucleotide and yielding an increase in fluorescence (**Figure 5.1A**). The Li group demonstrated that the beacon responded quickly ($T_{1/2}$ = 1.2 min) at 37° C, but was considerably slower at room temperature.

We have now used this displacement model to design additional aptamer beacon constructs. However, we have altered the model so that the fluorescent label is incorporated directly at the 5' end of the aptamer. In this way, only one antisense oligonucleotide is used to form the quenched, bipartite complex, and the release of this single inhibitor results immediately in a fluorescent signal (**Figure 5.1B**, **1C**). Since the only design variables for the two-piece constructs are the number and position of base-pairs between the aptamer and the antisense quencher oligonucleotide their performance can be more readily predicted, engineered, and optimized.

Optimizing quencher length for responsivity at room temperature

In order to better understand how aptamer beacon signaling and kinetics might be optimized we attempted to develop a sensor that could be used conveniently at room temperature in real-time. We first generated a series of aptamer beacons in which the number and positions of base-pairs between the antisense quencher oligonucleotide and the aptamer beacon were changed (**Figure 5.2A**; Apt4). Apt4 was synthesized with a fluorescein at its 5' end and a series of increasingly shorter antisense oligonucleotides were synthesized with DABCYL at their 3' ends. It was hypothesized that the shorterer quencher oligonucleotides would form less stable complexes, and thus might activate more quickly at room temperature.

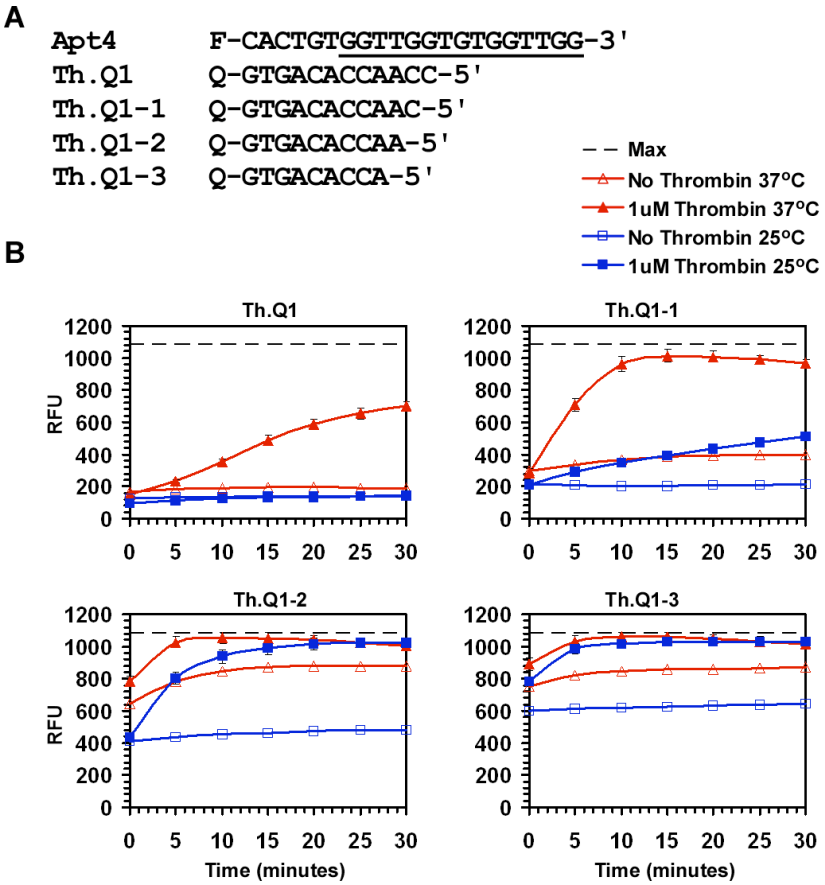


Figure 5.2. Activation of Apt4. A.) Apt4 and quencher sequences. F denotes fluorescein moiety, Q denotes DABCYL moiety. B.) Signal in relative fluorescent units (RFU) of Apt4 over 30 minutes at 26°C or 37°C in the absence (open marker) or presence of 1uM Thrombin (closed marker). Apt4 concentration was 100nM while quencher oligos were 200nM final 1X signaling buffer. The black dotted line represents Apt4 without quencher. Error bars represent standard deviation from triplicate samples.

The aptamer beacons were constructed by heating the oligonucleotides together then cooling them to room temperature; the initial fluorescent signals were stable prior to the addition of thrombin. Different beacons were quenched to different extents. The larger number of base-pairs gave lower backgrounds, with maximal quenching of fluorescence (8-fold) observed for Th.Q1 at 26°C. Upon the addition of 1uM (saturating) thrombin, fluorescence was monitored and the reaction was allowed to come to

equilibrium over 30 minutes at either 26° C or 37° C. As can be seen in **Figure 5.2B**, the smaller number of base-pairs generally gave larger final fluorescent signals.

By varying the effects of base-pairing on both background and final signal were observed, the interactions between the anti-thrombin aptamer and its antisense quencher could be poised to yield different extents and kinetics of activation. Th.Q1 is analogous to QDNA6 (Nutiu and Li 2003), and showed 3.7-fold activation in our hands at 37° C after 30 minutes, but very little activation at 26° C. These results differed from those originally reported. Nutiu and Li (2003) found ~7 fold activation of the beacon in less than 3 minutes at 37° C, and ~8-fold activation at 25° C after 30 minutes. The difference in results may be due to the fact that our beacon is a two-piece rather than a three-piece construct. Our other two-piece designed beacons showed smaller activations: a 2.4-fold increase in signal over background was seen with Th.Q1-1 at 37° C, and a 3.0-fold increase with Th.Q1-2 at 26° C.

As the extent of activation decreased, the rate of activation increased, as predicted (**Table 5.1**). Th.Q1-2 showed the fastest activation at room temperature ($T_{1/2}$ = 3.6 min), while Th.Q1-1 showed the fastest overall activation ($T_{1/2}$ = 3.4 min at 37° C).

Optimizing aptamer length for responsivity at room temperature

Having found a two-piece aptamer beacon pair that was capable of activation within 30 minutes at room temperature, we began to modify other design variables to see what effects they might have on the extent and kinetics of activation. Instead of modifying the length of the antisense quencher oligonucleotide, we instead generated a number of shorter aptamer beacons (Apt6, Apt7 and Apt8; **Figure 5.3A**). We assayed the new aptamer beacons with both the antisense quencher oligonucleotide Th.Q1, which was previously shown to fully quench, and ThQ1-2, which provided the best activation and one of the best response times. As previously, decreasing the length and strength of

hybridization led to greater background at room temperature (**Figure 5.3B**). Th.Q1-2 did not quench either Apt7 or Apt8, and neither antisense quencher oligonucleotide proved capable of quenching Apt6 (data not shown).

Apt4 with Th.Q1-2 and Apt8 with Th.Q1 displayed similar activation rates at room temperature (**Table 5.1**; $T_{1/2}$ values of 3.6 and 3.3 minutes, respectively). The observed inverse correlation between extents and rates of activation was found to be true over numerous aptamer beacons **(Figure 5.6A)**. In general, the fastest aptamer beacons were those that formed stable but weak complexes, and thus were poised to be perturbed by the addition of ligand. It appears as though a complex that has a $\Delta G°$ value of ca. -11 to -12 kcal / mole leads to improved kinetic performance. Indeed, there appears to be a quite steep kinetic cost to hybridization energies greater than -12 kcal / mole.

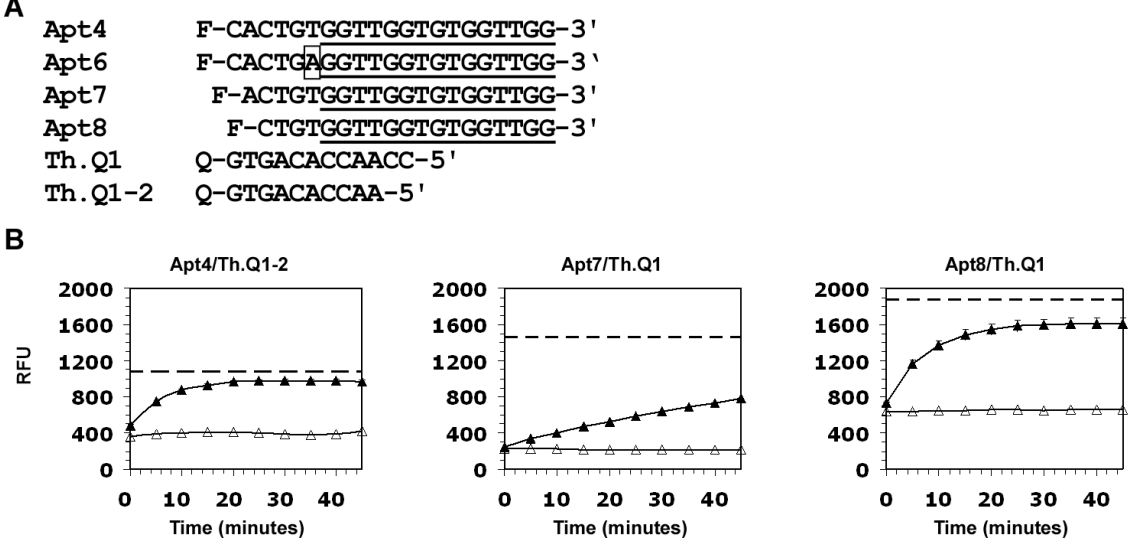


Figure 5.3. Effects of Aptamer Stem Minimization. A.) Aptamer and quencher sequences. F denotes fluorescein moiety, Q denotes DABCYL moiety. The boxed "A" in Apt6 represents a mispairing with the quencher oligo. B.) Signal in relative fluorescent units (RFU) of Apt4 and Th.Q1-2, Apt7 and Th.Q1, and Apt8 and Th.Q1 over 30 minutes at 26°C in the absence (open marker) or presence of 1uM Thrombin (closed marker). Aptamer concentration was 100nM while quencher oligos were 200nM in 1X signaling buffer. The black dotted line represents maximum signal for each aptamer (in the absence of quencher). Error bars represent one standard deviation taken from triplicate samples.

Effects of quencher type and position on aptamer beacon responsivities

Variable quenching efficiencies due to either contact quenching or FRET have previously been observed and measured with other hybridizing oligonucleotides (Dietrich, Buschmann et al. 2002; Marras, Kramer et al. 2002). We therefore wanted to see to whether quencher:fluor interactions also impacted the activation kinetics of our fastest aptamer beacons. We tested both Apt4 and Apt8 with two different antisense oligonucleotides, one which led to Q and F being in close proximity (Apt4/Th.Q1 or Apt8/Th.Q8) and another in which Q was separated by a two nucleotide spacer from F (Apt4/Th.Q8 or Apt8/Th.Q1; **Figure 5.4A**). The Apt4/Th.Q1 complexes proved to be too stable, and did not activate at room temperature. For the other aptamer beacons, we also attempted to determine whether there were performance differences between commonly used quencher moieties, such as BHQ-1, dabcyl and Eclipse Quencher.

The extents of activation seen with the different the quenchers were generally similar irrespective of the antisense quencher oligonucleotide or aptamer beacon complex. However, it appeared that aptamer beacons that displaced quencher moieties away from paired strands (Apt4/Th.Q1-2 and Apt 8/Th.Q8) had somewhat lower extents of activation but faster kinetics. This may imply that the quencher moieties interact with the paired strands. This interpretation is consistent with previous findings, in that duplexes labeled with quenchers showed from 2° to 10° C greater melting temperatures compared with unlabeled duplexes (Marras et al., 2002).

A

```
Apt4      F-CACTGTGGTTGGTGTGGTTGG-3'
Apt8        F-CTGTGGTTGGTGTGGTTGG-3'
Th.Q1     Q-GTGACACCAACC-5'
Th.Q8       Q-GACACCAACC-5'
```

B

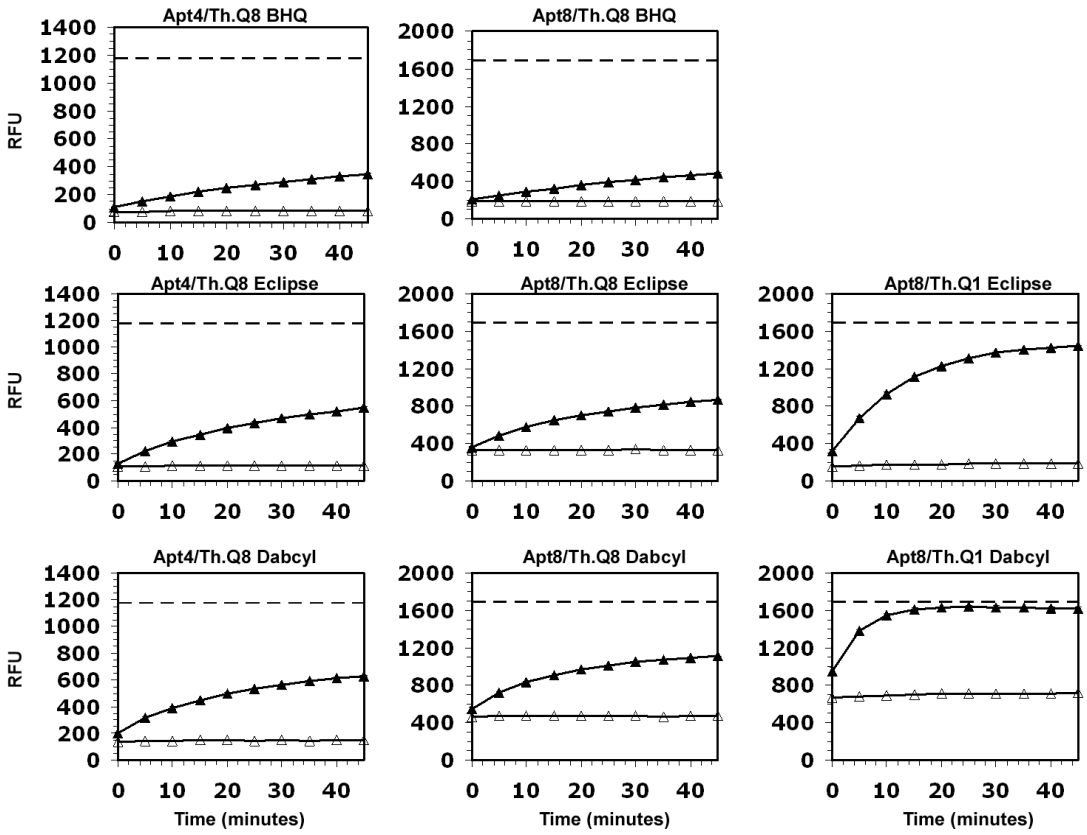


Figure 5.4. Varying quencher and position. A.) Aptamer and quencher sequences. F denotes fluorescein moiety, Q denotes either BHQ-1, Dabcyl, or Eclipse Quencher moieties as specified in B.) Signal in relative fluorescent units (RFU) of Apt8 or Apt4 with either the Th.Q1 or ThQ8 quencher oligos and three different quencher moieties over 45 minutes at 26°C in the absence (open marker) or presence of 1uM Thrombin (closed marker). Aptamer concentration was 100nM while quencher oligos were 200nM in 1X signaling buffer. The black dotted line represents maximum signal for each aptamer (in the absence of quencher). Error bars represent one standard deviation taken from triplicate samples.

Interestingly, quencher:strand interactions seem to affect kinetics much more strongly than activation. For example, when comparing Th.Q8 (quencher displaced) with Th.Q1 (quencher overlap) on Apt8, the extents of activation were almost all the same

(**Table 5.1**), yet the different quenchers yielded very different rates of activation. The aptamer beacons that incorporated Dabcyl had two- to three-times faster ($T_{1/2}$) responses, while those with BHQ-1 displayed the slowest responses. Previous comparisons between these quenchers also showed that Dabcyl had a lower impact on melting temperature of duplexes than BHQ-1 (Marras et al., 2002). The slow BHQ-1-derivatized beacons also displayed the highest level of quenching (up to 94%, very similar to previously published value of 93% from Marras et al., 2002). It is possible that quencher interactions with the duplex not only stabilize the complexes but affect the 'breathing' of the strands (Mergny and Maurizot 2001; Hong, Harbron et al. 2003; Voss, Meyer et al. 2004; De Cian, Guittat et al. 2007) that may be necessary for faster switching between conformations.

Designing ligand accessibility

Nutiu and Li (2003) also designed a tripartite aptamer beacon for sensing ATP, but found that their original design did not perform well at room temperature. They hypothesized that this was because the quencher oligonucleotide "occupied" the aptamer core sequence, presumably preventing nascent folding and ligand interactions. As a test of this hypothesis, they constructed a second ATP-sensing aptamer beacon with additional sequence at its 5' end, outside of the aptamer core. The extended binding site interacted with a quencher oligonucleotide with the same thermodynamic parameters as with the failed construct. However, this new construct yielded activation at room temperature with ATP.

We did not explicitly take into account an "occupation" parameter in our initial designs. However, in Apt8/Th.Q1 the antisense oligonucleotide occupies the core quadruplex much more than in Atp4/Th.Q1-2, yet the two aptamer beacons have similar activation extents and kinetics. That said, the core quadruplex is occupied to some extent by the antisense quencher oligonucleotide in both beacons.

Rather than merely extend previous constructs, we chose to use a similar anti-thrombin aptamer that was already extended beyond the core quadruplex structure. Tasset et al. (1997) had previously selected an anti-thrombin aptamer that once again formed a quadruplex structure, but also contained an extended stem (aptamer 60-18[29]; Tasset, Kubik et al. 1997; Figure 5.1C). Although the original (Bock) aptamer and the extended (Tasset) aptamer both form stacked G-quartets they contain different connecting loops and are thought to preferentially bind to different sites on thrombin. The Bock aptamer is thought to bind to thrombin's fibrinogen recognition exosite, while the Tasset aptamer binds to the heparin exosite as discussed above.

We designed two-piece aptamer beacons based on the Tasset aptamer. Since the additional base-pairs formed by the aptamer meant that the structure would be more difficult to denature, we used a series of antisense quencher oligonucleotides that hybridized with predicted ΔG° values of -11.3 kcal / mole to -15.5 kcal / mole (**Figure 5.1C**; **Table 5.1**). In order to explicitly test whether occupancy was a factor in aptamer beacon kinetic performance, an additional six base 'handle' (TCAGCA) was added at the 5' end; this moved the antisense quencher oligonucleotides completely outside of the quandruplex forming region of the aptamer (**Figure 5.5A**).

A

```
AptH       F-TCAGCAAGTCCGTGGTAGGGCAGGTTGGGGTGACT-3'
Th.Q5      Q-AGTCGTTCAGGC-3'
Th.Q5-1    Q-AGTCGTTCAGG-3'
Th.Q5-2    Q-AGTCGTTCAG-3'
Th.Q3-1     Q-GTCGTTCAGGC-3'
Th.Q3-2      Q-TCGTTCAGGC-3'
```

B

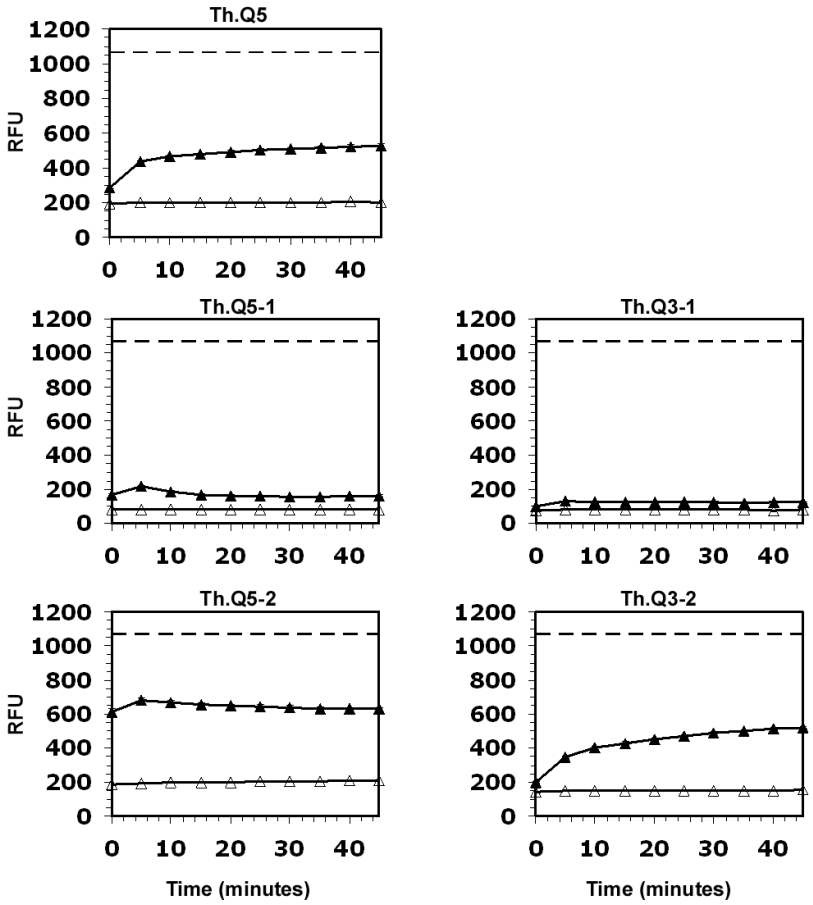


Figure 5.5. AptH aptamer Activation. A.) Aptamer and quencher sequences. F denotes fluorescein moiety, Q denotes DABCYL moiety. The boxed sequence forms a duplex stem while the grey represents an internal bulge characteristic of this aptamer. B.) Signal in relative fluorescent units (RFU) of the sequences in (A) over 45 minutes at 26°C in the absence (open marker) or presence of 1uM Thrombin (closed marker). Aptamer concentration was 100nM while quencher oligos were 200nM in 1X signaling buffer. The black dotted line represents maximum signal for each aptamer (in the absence of quencher). Error bars represent one standard deviation taken from triplicate samples.

As before, the least stable complexes (AptH/Th.Q3-2 and AptH/Th.Q5-2) gave the fastest activation (**Figure 5.5B**). It should be noted that even when fully activated these aptamer beacons showed only ca. 60% of the fluorescence seen in the absence of the quencher. We speculated that this might be because of protein-mediated quenching, but assays with protein and no quencher oligonucleotide showed full fluorescence. Thus, it seems likely that an equilibrium between bound protein and bound quencher has been established. This is interesting because it implies that these biosensors may have reversible performance.

More stable complexes (AptH/Th.Q3-1 and AptH/Th.Q5-1) gave little or no activation, with the exception of the most stable, AptH/Th.Q5. Interestingly, while AptH/Th.Q5 is predicted to be the most stable complex, it gave a higher background reading than the aptamer beacons that showed virtually no activation. We believe that AptH/Th.Q5 was not fully quenched because of a competition between binding to the aptamer and a fortuitous dimerization of the antisense quencher oligonucleotide itself, and that this additional equilibria also led to the unexpected activation by thrombin.

Even though we predicted that a stronger interaction with the antisense quencher oligonucleotide would be necessary for aptamer beacon complex formation, it was again an antisense quencher oligonucleotide that bound with a ΔG° between -11 kcal / mole and -12 kcal / mole that gave the best performance (AptH/Th.Q5-2; -11.3 kcal / mole). This aptamer beacon had remarkable kinetic responsivity, and was almost fully activated within 1 minute.

The extraordinary speed of AptH/Th.Q5-2 may have been because the quadruplex structure that interacts with thrombin could pre-form more readily than in previous (Apt4, Apt8) constructs. In this mechanistic model, nascent interactions with the protein might assist in displacing the antisense quencher oligonucleotide. A similar model has been

proposed for the function of so-called antiswitches (Bayer and Smolke 2005), and is operative in nucleic acid logic circuits based on 'toehold' sequences (Kim, White et al. 2006; Seelig, Soloveichik et al. 2006).

Predicting the kinetics of aptamer beacons

Overall there was a strong correlation between the melting temperatures of the antisense quencher oligonucleotides and the speeds of response of the aptamer beacons (**Figure 5.6B,C**). This trend can of course be rationalized by suggesting that the antisense oligonucleotide must be released prior to forming the ligand-binding conformation, and thus that the speeds of the various aptamer beacons were determined largely by the off-rates of the antisense quencher oligonucleotides. If so, this would seem to argue against the relevance of the 'protein assist' model for AptH/Th.Q5-2 described above. Interestingly, though, this aptamer beacon does have an unusually high F/F_o value, falling well outside of the trend previously seen in **Figure 5.6A**. It may be that the simultaneous formation of the quadruplex and antisense duplex structures yields a low energy conformation that is only activated (albeit quickly) by saturating protein concentrations.

Given that the apparent stabilities of several of our aptamer beacon complexes differed from their predicted stabilities, based on both the observed background fluorescence and speeds of response, we decided to directly measure the T_m values of the aptamer beacons. To do this, beacons were prepared as in the activation assays, then ramped from 10°C to 95°C at a rate of ~0.1° C/sec and read on a real time PCR machine after each 1° C increase. A revised plot of measured T_m values versus $T_{1/2}$ showed a remarkable congruence (**Figure 5.6C**; r^2 = **0.59**), even when the thermodynamic effects attributed to the Eclipse Quencher were included. These results indicate that the kinetic responsivities of aptamer beacons can be largely predicted based on this single parameter.

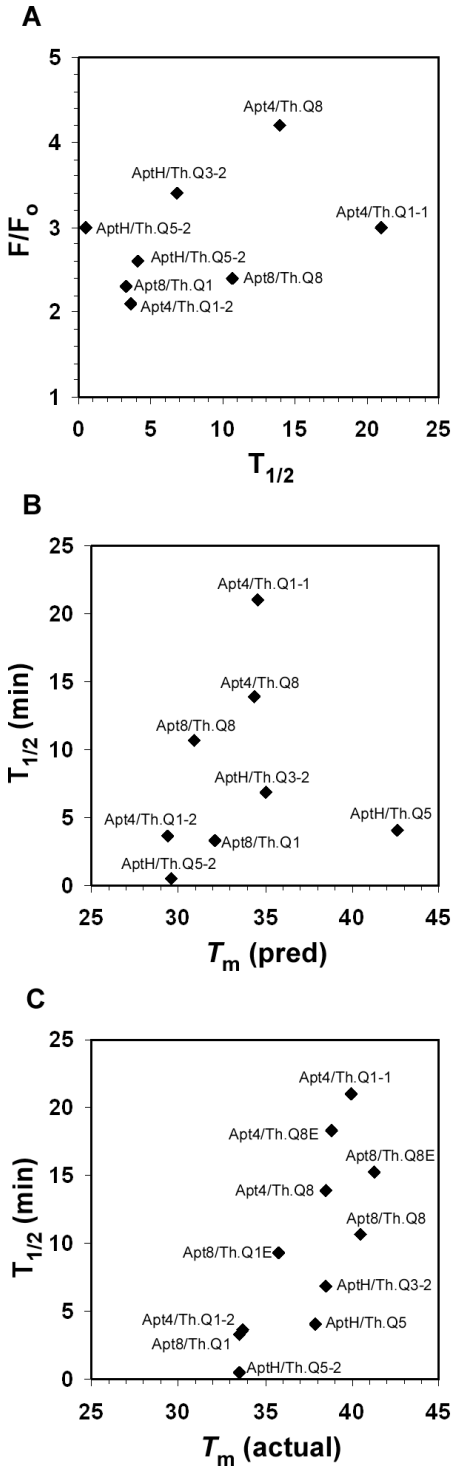


Figure 5.6. Comparison of thermodynamic and kinetic parameters from Table 5.1 for a subset of active aptamer beacons.

Name	Sequence	Predicted ΔG	Predicted T_m	Actual T_m	# bp with quencher	Fold Activation*	% Min, Max Fluorescence	$T_{1/2}$ (min)**	Rate (min^{-1})**
5F.Apt.4	CACTGTGGTTGGTGTGGTTGG								
Th.Q1	GTGACACCAACC	-14.6	39.8	45.6	12	1.2; 3.7	9%, 15%; 14%, 64%	> 30	< 0.023
Th.Q1-1	GTGACACCAAC	-12.8	34.6	40.0	11	3.0; 2.4	19%, 57%; 26%, 93%	21, 3.4	0.034, 0.21
Th.Q1-2	GTGACACCAA	-11.2	29.4	33.7	10	2.1; 1.1	40%, 94%; 72%, 97%	3.6	0.19
Th.Q1-3	GTGACACCA	-9.9	23.7	29.3	9	1.6; 1.2	72%, 95%; 82%, 97%	-	-
Th.Q8 (Dabcyl)		-12.7	34.4	38.5	10	4.2	12%, 54%	13.9	0.050
Th.Q8 (Eclipse)	GACACCAACC	-12.7	34.4	38.8	10	4.6	9%, 46%	18.3	0.038
Th.Q8 (BHQ)		-12.7	34.4	40.9	10	4.1	6%, 30%	>30	< 0.023
5F.Apt.6	CACTGAGGTTGGTGTGGTTGG								
Th.Q1	GTGACACCAACC	-10.1	24.6		12			-	-
5F.Apt.7	ACTGTGGTTGGTGTGGTTGG								
Th.Q1	GTGACACCAACC	-13.4	36.9	39.0	11	3.6	17%, 53%	> 30	< 0.023
5F.Apt.8	CTGTGGTTGGTGTGGTTGG								
Th.Q1 (Dabcyl)	GTGACACCAACC	-12.1	32.1	33.5	10	2.3	39%, 97%	3.3	0.21
Th.Q1 (Eclipse)		-12.1	32.1	35.8	10	7.7	9%, 85%	9.3	0.074
Th.Q8 (Dabcyl)		-11.7	30.9	40.5	10	2.4	28%, 66%	10.6	0.065
Th.Q8 (Eclipse)	GACACCAACC	-11.7	30.9	41.3	10	2.6	19%, 51%	15.2	0.046
Th.Q8 (BHQ)		-11.7	30.9	45.1	10	2.6	11%, 29%	> 30	< 0.023
5F.AptH	TCAGCAAGTCCGTGGTAGGGCAGGTTGGGGTGACT								
Th.Q5	AGTCGTTCAGGC	-15.5	42.6	37.9	12	2.6	18%, 49%	4.1	0.17
Th.Q5-1	AGTCGTTCAGG	-12.9	35.0	42.7	11	2.0	7%, 15%	-	-
Th.Q5-2	AGTCGTTCAG	-11.3	29.6	33.5	10	3.0	17%, 64%	< 1	> 0.69
Th.Q3-1	GTCGTTCAGGC	-14.6	40.0	46.3	11	1.5	7%, 12%	-	-
Th.Q3-2	TCGTTCAGGC	-12.7	35.0	38.5	10	3.4	13%, 49%	6.8	0.1014

Table 5.1. Thermodynamic profiles. The melting temperature is provided in °C, free energy is in kcal/mol. *Blue font color corresponds to data taken at 26° C while red font corresponds to 37°C ** $T_{1/2}$ and rate constants only reported where activation is over 2 fold and the R^2 value of the curve that fits the data to the equation $T_{1/2}$=-ln(y)/rate is > 0.95 as calculated by SigmaPlot v9.0. This table reports % minimum fluorescence which can be converted to quenching efficiency by subtracting from 100% as done in Marras *et al* (2002).

CONCLUSIONS AND FUTURE DIRECTIONS

The work in this entire dissertation is not in a field by itself. Much of the work builds on previous aptamer based biosensor developments in both signaling and detection methods. We have combined rational design with computer-assisted prediction of nucleic acid conformational changes to mimic aptamer selections *in silico*. These methods however require a fair amount of in depth knowledge that cannot be programmed and predicted by a computer currently.

While we have produced a number of functional aptamer biosensors from computationally derived and selected pools, we have not exhausted or even touched on the potential functional sequences present in these pools. We have not provided a statistically relevant sample set to assume the sequences tested are the only sequences capable of switching. In fact, we have not tested a sufficient number of sequences from the subset of the pool rejected by our computational sieving mechanisms to prove those in fact do not work. Many of these limitations are due to synthesis costs for creating and purifying many different RNAs, single and dual labeled probes. Therefore, for future systems, we believe computational optimization, the use of a cheap ligand, and modeling DNA aptamers that display tight binding to their cognate analyte and can be easily synthesized would be beneficial. The designs should be able to prototype many aptamer sequences with a small or limited number of fluorescent beacons or alternate detection mechanism.

In all of our methods, we have only calculated and compared structural thermodynamics to elucidate potential predictive parameters. It is apparent that the use of thermodynamic minimization dynamic programming software such as Mfold and ViennaRNA may limit the ability of our programs to accurately identify functional switching structures which are also affected by kinetic limitations. We could in the future use simulated annealing or kinetic information to predict biosensors that switch faster and with reduced backgrounds.

The Freedonia Group reports that the in vitro diagnostic market represented $15.2 billion in US market in 2006. While the use of aptamers in biomedical research is vast, there are currently no examples of commercial aptamer based biosensors development due to patenting issues. Much of the intellectual property for aptamer biosensors is held by SomaLogic and Gilead (formally NeXstar). In terms of aptamer

based therapeutic applications, much of the intellectual property is owned by Archemix. However, licensing is available for aptamer biosensors deemed commercially viable. Thus, much of the methods presented in this dissertation could be used to increase the efficiency of selecting functional species from smaller random pools and reduce the time and cost to generate functional aptamer based biosensors making them more commercially attractive.

ACKNOWLEDGEMENTS

We would like to thank Eun Jeong Cho for her helpful discussions on the various thrombin aptamers, biosensor design, and her help with the design of AptH. We would also like to thank Shawn Piasecki for her help in data collection.

MATERIALS AND METHODS

Synthesis and Purification

All materials were synthesized as in (Hall, Hesselberth et al. 2007). Breifly 5' 6-FAM moieties were introduced by 5' phosphoramidite chemistry, while 3' quenchers were incorporated through controlled pore glass (CPG) solid supports (Glen research, Sterling, VA) during synthesis. The quenchers used include dabcyl (4-(4-dimethylaminophenylazo)benzoic acid), Eclipse Quencher, and Black Hole Quencher 1 (BHQ-1). After deprotection samples were lyophilized and purified by HPLC (Nutiu and Li 2003)with the following changes: column was an XTerra MS C18 Column with dimentions of 4.6mm x 30mm and 2.5um bead diameter. The two solvents were 0.1M TEAA and 100% acetonitrile on a nonlinear gradient from 5% to 35% acetonitrile over 30 minutes at a flow rate of 3 mL/min. We collected 2/3 of the main peak width showing absorbtion at both 260nm and 495nm. The purified DNA was lyophilized and

resuspended in 100ul diH_2O. Strand concentrations were detected by UV spectroscopy at 260nm on a Nanodrop ND-1000 (Wilmington, DE) using extinction coefficients and adjusted to 10uM for each oligonucleotide sequence.

Kinetic Assays

Assays were performed with 100nM fluorescent oligo and 200nM quencher oligo in 50uL reactions. Briefly, both were mixed in signaling buffer (20mM Tris pH 8.3, 5mM KCl, 1mM $MgCl_2$ (Nutiu and Li 2003)), heat denatured at 90° C for 3 minutes, cooled to 25° C at 0.2° C/sec. They were allowed to further thermally equlibrate at room temperature (26° C) for 5 to 15 minutes. Assays were carried out in Corning 96 well black opaque ½ area well plates (Corning, NY). Maximum fluorescence was measured in the absence of quencher oligonucleotide for each assay. Aptamers were assayed for activation in the presence or absence of 1uM thrombin (Haematologic Technologies, Essex jct., VT) on a BioTek Synergy HT (Winooski, VT) fluorescent plate reader at between 25.7° C and 26.8° C with a gain of either 48 or 50 such that a buffer blank yielded a signal of 26 RFU. There was a lag of between 30 seconds and 1 minute after thrombin addition and time point 0. Due to this lag and thermal accuracy near room temperature (26° C) with the Synergy HT, thermally controlled kinetic assays were monitored in real time on a PTI fluorimeter (Birmingham, NJ) and used to calculate rate constants with the equation $T_{1/2} = -\ln(0.5)/\text{rate}$ for the fastest constructs.

Thermodynamic Predictions

The DINAmelt 2-state Hybridization Server (http://frontend.bioinfo.rpi.edu/applications/hybrid/twostate.php) was used to predict melting profiles of each nucleic acid pair (Table 1; Mathews, Sabina et al. 1999; Zuker 2003). The following parameters were modified from the default settings: Energy Rules:

DNA at 26° C; Salts: 10mM Na^+ (The lowest allowed although we have little Na+ in our reactions) 1mM Mg^{++}; Strand Concentration: 0.1uM. There is no way to calculate the effects of strand concentration through this web interface.

Fluorescent Melting Assays

Melts were performed on an Applied Biosystems 7300 Real Time PCR System (place) in skirted 96 well PCR plates from the same company using the dissociation method. Assays were set up identical to the kinetic assays. Samples were heated to 95°C for 15 seconds, then cooled to 10°C within the method. Temperature ramps were performed from 10° C to 95° C at a rate of ~0.1° C/sec then read after each 1° C increase. Samples were repeated in triplicate for three complete transition curves and normalized to the minimum and maximum fluorescence. The melting temperature (T_m) was calculated as the peak of a curve fit through the first derivative of the normalized sigmoidal melt curve (Mergny and Lacroix 2003).

References

(2006). The Merk index. Whitehouse Station, Merck & Co., Inc.

Andronescu, M., A. P. Fejes, F. Hutter, H. H. Hoos and A. Condon (2004). "A new algorithm for RNA secondary structure design." J Mol Biol **336**(3): 607-24.

Andronescu, M., Z. C. Zhang and A. Condon (2005). "Secondary structure prediction of interacting RNA molecules." J Mol Biol **345**(5): 987-1001.

Avihoo, A. and D. Barash (2006). "Shape similarity measures for the design of small RNA switches." J Biomol Struct Dyn **24**(1): 17-24.

Avihoo, A., I. Gabdank, M. Shapira and D. Barash (2007). "In silico design of small RNA switches." IEEE Trans Nanobioscience **6**(1): 4-11.

Bayer, T. S. and C. D. Smolke (2005). "Programmable ligand-controlled riboregulators of eukaryotic gene expression." Nat Biotechnol **23**(3): 337-43.

Bell, S. D., J. M. Denu, J. E. Dixon and A. D. Ellington (1998). "RNA molecules that bind to and inhibit the active site of a tyrosine phosphatase." J Biol Chem **273**(23): 14309-14.

Bier, F. F. and J. P. Furste (1997). "Nucleic acid based sensors." Exs **80**: 97-120.

Bock, L. C., L. C. Griffin, J. A. Latham, E. H. Vermaas and J. J. Toole (1992). "Selection of single-stranded DNA molecules that bind and inhibit human thrombin." Nature **355**(6360): 564-6.

Bonnet, G., S. Tyagi, A. Libchaber and F. R. Kramer (1999). "Thermodynamic basis of the enhanced specificity of structured DNA probes." Proc Natl Acad Sci U S A **96**(11): 6171-6.

Breaker, R. R. (2002). "Engineered allosteric ribozymes as biosensor components." Curr Opin Biotechnol **13**(1): 31-9.

Brion, P. and E. Westhof (1997). "Hierarchy and dynamics of RNA folding." Annu Rev Biophys Biomol Struct **26**: 113-37.

Bunka, D. H. and P. G. Stockley (2006). "Aptamers come of age - at last." Nat Rev Microbiol **4**(8): 588-96.

Burgstaller, P., M. Kochoyan and M. Famulok (1995). "Structural probing and damage selection of citrulline- and arginine-specific RNA aptamers identify base positions required for binding." Nucleic Acids Res **23**(23): 4769-76.

Chen, X. Y., T. M. Link and V. L. Schramm (1998). "Ricin A-chain: kinetics, mechanism, and RNA stem-loop inhibitors." Biochemistry **37**(33): 11605-13.

Cho, E. J., J. R. Collett, A. E. Szafranska and A. D. Ellington (2006). "Optimization of aptamer microarray technology for multiple protein targets." Anal Chim Acta **564**(1): 82-90.

Cho, E. J., J.-W. Lee, M. Rajendran and A. D. Ellington (In Press). Nucleic acids for reagentless biosensors. Amsterdam, Neth, Elsevier Science B.V.

Cho, E. J., M. Rajendran and A. D. Ellington (In Press). Aptamers as emerging probes for macromilecular sensing. Topics in Fluorescence Spectroscopy. J. R. Lakowicz. New York, Plenum Press. **9**.

Collett, J. R., E. J. Cho and A. D. Ellington (2005). "Production and processing of aptamer microarrays." Methods **37**(1): 4-15.

Danilova, L. V., D. D. Pervouchine, A. V. Favorov and A. A. Mironov (2006). "RNAKinetics: a web server that models secondary structure kinetics of an elongating RNA." J Bioinform Comput Biol **4**(2): 589-96.

De Cian, A., L. Guittat, M. Kaiser, B. Sacca, S. Amrane, A. Bourdoncle, P. Alberti, M. P. Teulade-Fichou, L. Lacroix and J. L. Mergny (2007). "Fluorescence-based melting assays for studying quadruplex ligands." Methods **42**(2): 183-95.

Deisingh, A. K. (2006). "Aptamer-based biosensors: biomedical applications." Handb Exp Pharmacol(173): 341-57.

Dietrich, A., V. Buschmann, C. Muller and M. Sauer (2002). "Fluorescence resonance energy transfer (FRET) and competing processes in donor-acceptor substituted DNA strands: a comparative study of ensemble and single-molecule data." J Biotechnol **82**(3): 211-31.

Ding, Y. (2006). "Statistical and Bayesian approaches to RNA secondary structure prediction." Rna **12**(3): 323-31.

Ding, Y. and C. E. Lawrence (2003). "A statistical sampling algorithm for RNA secondary structure prediction." Nucleic Acids Res **31**(24): 7280-301.

Dirks, R. M., J. S. Bois, J. M. Schaeffer, E. Winfree and N. A. Pierce (2007). "Thermodynamic Analysis of Interacting Nucleic Acid Strands." SIAM Review **49**(1): 65-88.

Dirks, R. M., M. Lin, E. Winfree and N. A. Pierce (2004). "Paradigms for computational nucleic acid design." Nucleic Acids Res **32**(4): 1392-403.

Dirks, R. M. and N. A. Pierce (2004). "Triggered amplification by hybridization chain reaction." Proc Natl Acad Sci U S A **101**(43): 15275-8.

Doshi, K. J., J. J. Cannone, C. W. Cobaugh and R. R. Gutell (2004). "Evaluation of the suitability of free-energy minimization using nearest-neighbor energy parameters for RNA secondary structure prediction." BMC Bioinformatics **5**: 105.

Dowell, R. D. and S. R. Eddy (2004). "Evaluation of several lightweight stochastic context-free grammars for RNA secondary structure prediction." BMC Bioinformatics **5**: 71.

Drolet, D. W., L. Moon-McDermott and T. S. Romig (1996). "An enzyme-linked oligonucleotide assay." Nat Biotechnol **14**(8): 1021-5.

Eddy, S. R. (2004). "How do RNA folding algorithms work?" Nat Biotechnol **22**(11): 1457-8.

Eddy, S. R. (2004). "What is dynamic programming?" Nat Biotechnol **22**(7): 909-10.

Edwards, T. E., D. J. Klein and A. R. Ferre-D'Amare (2007). "Riboswitches: small-molecule recognition by gene regulatory RNAs." Curr Opin Struct Biol **17**(3): 273-9.

Ellington, A. D. and J. W. Szostak (1990). "In vitro selection of RNA molecules that bind specific ligands." Nature **346**(6287): 818-22.

Fan, P., A. K. Suri, R. Fiala, D. Live and D. J. Patel (1996). "Molecular recognition in the FMN-RNA aptamer complex." J Mol Biol **258**(3): 480-500.

Fang, X., A. Sen, M. Vicens and W. Tan (2003). "Synthetic DNA aptamers to detect protein molecular variants in a high-throughput fluorescence quenching assay." Chembiochem **4**(9): 829-34.

Feigon, J., T. Dieckmann and F. W. Smith (1996). "Aptamer structures from A to ζ." Chem. Biol. **3**(8): 611-617.

Flamm, C., W. Fontana, I. L. Hofacker and P. Schuster (2000). "RNA folding at elementary step resolution." Rna **6**(3): 325-38.

Flamm, C., I. L. Hofacker, S. Maurer-Stroh, P. F. Stadler and M. Zehl (2001). "Design of multistable RNA molecules." Rna **7**(2): 254-65.

Frauendorf, C. and A. Jaschke (2001). "Detection of small organic analytes by fluorescing molecular switches." Bioorg Med Chem **9**(10): 2521-4.

Fredriksson, S., M. Gullberg, J. Jarvius, C. Olsson, K. Pietras, S. M. Gustafsdottir, A. Ostman and U. Landegren (2002). "Protein detection using proximity-dependent DNA ligation assays." Nat Biotechnol **20**(5): 473-7.

Friedrichs, E. and F. C. Simmel (2007). "Controlling DNA polymerization with a switchable aptamer." Chembiochem **8**(14): 1662-6.

Gautheret, D., D. Konings and R. R. Gutell (1994). "A major family of motifs involving G.A mismatches in ribosomal RNA." J Mol Biol **242**(1): 1-8.

Geiger, A., P. Burgstaller, H. von der Eltz, A. Roeder and M. Famulok (1996). "RNA aptamers that bind L-arginine with sub-micromolar dissociation constants and high enantioselectivity." Nucleic Acids Res **24**(6): 1029-36.

Giegerich, R., D. Haase and M. Rehmsmeier (1999). "Prediction and visualization of structural switches in RNA." Pac Symp Biocomput: 126-37.

Gouda, H., I. D. Kuntz, D. A. Case and P. A. Kollman (2003). "Free energy calculations for theophylline binding to an RNA aptamer: Comparison of MM-PBSA and thermodynamic integration methods." Biopolymers **68**(1): 16-34.

Green, S. J., D. Lubrich and A. J. Turberfield (2006). "DNA hairpins: fuel for autonomous DNA devices." Biophys J **91**(8): 2966-75.

Gutell, R. R., J. C. Lee and J. J. Cannone (2002). "The accuracy of ribosomal RNA comparative structure models." Curr Opin Struct Biol **12**(3): 301-10.

Haes, A. J., B. C. Giordano and G. E. Collins (2006). "Aptamer-based detection and quantitative analysis of ricin using affinity probe capillary electrophoresis." Anal Chem **78**(11): 3758-64.

Hall, B., J. R. Hesselberth and A. D. Ellington (2007). "Computational selection of nucleic acid biosensors via a slip structure model." Biosens Bioelectron **22**(9-10): 1939-47.

Hamaguchi, N., A. Ellington and M. Stanton (2001). "Aptamer beacons for the direct detection of proteins." Anal Biochem **294**(2): 126-31.

Hermann, T. and D. J. Patel (2000). "Adaptive recognition by nucleic acid aptamers." Science **287**(5454): 820-5.

Hesselberth, J. R., D. Miller, J. Robertus and A. D. Ellington (2000). "In vitro selection of RNA molecules that inhibit the activity of ricin A-chain." J Biol Chem **275**(7): 4937-42.

Heyduk, E. and T. Heyduk (2005). "Nucleic acid-based fluorescence sensors for detecting proteins." Anal Chem **77**(4): 1147-56.

Hianik, T., V. Ostatna, M. Sonlajtnerova and I. Grman (2007). "Influence of ionic strength, pH and aptamer configuration for binding affinity to thrombin." Bioelectrochemistry **70**(1): 127-33.

Ho, H. A. and M. Leclerc (2004). "Optical sensors based on hybrid aptamer/conjugated polymer complexes." J Am Chem Soc **126**(5): 1384-7.

Hofacker, I. L. (2003). "Vienna RNA secondary structure server." Nucleic Acids Res **31**(13): 3429-31.

Hofacker, I. L., W. Fontana, P. F. Stadler, L. S. Bonhoeffer, M. Tacker and P. Schuster (1994). "Fast folding and comparison of RNA secondary structures." Monatsh. Chem. **125**(2): 167-88.

Hong, M. K., E. J. Harbron, D. B. O'Connor, J. Guo, P. F. Barbara, J. G. Levin and K. Musier-Forsyth (2003). "Nucleic acid conformational changes essential for HIV-1 nucleocapsid protein-mediated inhibition of self-priming in minus-strand transfer." J Mol Biol **325**(1): 1-10.

Huang, F., G. Wang, T. Coleman and N. Li (2003). "Synthesis of adenosine derivatives as transcription initiators and preparation of 5' fluorescein- and biotin-labeled RNA through one-step in vitro transcription." Rna **9**(12): 1562-70.

Huizenga, D. E. and J. W. Szostak (1995). "A DNA aptamer that binds adenosine and ATP." Biochemistry **34**(2): 656-65.

Jayasena, S. D. (1999). "Aptamers: an emerging class of molecules that rival antibodies in diagnostics." Clin Chem **45**(9): 1628-50.

Jellinek, D., L. S. Green, C. Bell and N. Janjic (1994). "Inhibition of receptor binding by high-affinity RNA ligands to vascular endothelial growth factor." Biochemistry **33**(34): 10450-6.

Jenison, R. D., S. C. Gill, A. Pardi and B. Polisky (1994). "High-resolution molecular discrimination by RNA." Science **263**(5152): 1425-9.

Jhaveri, S., M. Rajendran and A. D. Ellington (2000). "In vitro selection of signaling aptamers." Nat Biotechnol **18**(12): 1293-7.

Jhaveri, S. D., R. Kirby, R. Conrad, E. J. Maglott, M. Bowser, R. T. Kennedy, G. Glick and A. D. Ellington (2000). "Designed Signaling Aptamers that Transduce Molecular Recognition to Changes in Fluorescence Intensity." J. Am. Chem. Soc. **122**(11): 2469-2473.

Jiang, Y., X. Fang and C. Bai (2004). "Signaling aptamer/protein binding by a molecular light switch complex." Anal Chem **76**(17): 5230-5.

Jose, A. M., G. A. Soukup and R. R. Breaker (2001). "Cooperative binding of effectors by an allosteric ribozyme." Nucleic Acids Res **29**(7): 1631-7.

Jucker, F. M., R. M. Phillips, S. A. McCallum and A. Pardi (2003). "Role of a heterogeneous free state in the formation of a specific RNA-theophylline complex." Biochemistry **42**(9): 2560-7.

Katilius, E., Z. Katiliene and N. W. Woodbury (2006). "Signaling aptamers created using fluorescent nucleotide analogues." Anal Chem **78**(18): 6484-9.

Kim, D. S., V. Gusti, S. G. Pillai and R. K. Gaur (2005). "An artificial riboswitch for controlling pre-mRNA splicing." Rna **11**(11): 1667-77.

Kim, J., K. S. White and E. Winfree (2006). "Construction of an in vitro bistable circuit from synthetic transcriptional switches." Mol Syst Biol **2**: 68.

Kirby, R., E. J. Cho, B. Gehrke, T. Bayer, Y. S. Park, D. P. Neikirk, J. T. McDevitt and A. D. Ellington (2004). "Aptamer-based sensor arrays for the detection and quantitation of proteins." Anal Chem **76**(14): 4066-75.

Kubik, M. F., A. W. Stephens, D. Schneider, R. A. Marlar and D. Tasset (1994). "High-affinity RNA ligands to human alpha-thrombin." Nucleic Acids Res **22**(13): 2619-26.

Lee, J. F., G. M. Stovall and A. D. Ellington (2006). "Aptamer therapeutics advance." Curr Opin Chem Biol **10**(3): 282-9.

Leontis, N. B., A. Lescoute and E. Westhof (2006). "The building blocks and motifs of RNA architecture." Curr Opin Struct Biol **16**(3): 279-87.

Li, J. J., X. Fang and W. Tan (2002). "Molecular aptamer beacons for real-time protein recognition." Biochem Biophys Res Commun **292**(1): 31-40.

Ligler, F. S., D. Phil and S. D. (2006). Biosensors for Detection of Bioterrorists Threats. Optical Chemical Sensors. F. Baldini, A. N. Chester, J. Homola and S. Martellucci, Springer Netherlands. **224:** 437-455.

Lord, J. M., M. R. Hartley and L. M. Roberts (1991). "Ribosome inactivating proteins of plants." Semin Cell Biol **2**(1): 15-22.

Macaya, R. F., P. Schultze, F. W. Smith, J. A. Roe and J. Feigon (1993). "Thrombin-binding DNA aptamer forms a unimolecular quadruplex structure in solution." Proc Natl Acad Sci U S A **90**(8): 3745-9.

Macaya, R. F., J. A. Waldron, B. A. Beutel, H. Gao, M. E. Joesten, M. Yang, R. Patel, A. H. Bertelsen and A. F. Cook (1995). "Structural and functional characterization of potent antithrombotic oligonucleotides possessing both quadruplex and duplex motifs." Biochemistry **34**(13): 4478-92.

Marras, S. A., F. R. Kramer and S. Tyagi (2002). "Efficiencies of fluorescence resonance energy transfer and contact-mediated quenching in oligonucleotide probes." Nucleic Acids Res **30**(21): e122.

Mathews, D. H. (2004). "Using an RNA secondary structure partition function to determine confidence in base pairs predicted by free energy minimization." Rna **10**(8): 1178-90.

Mathews, D. H. (2006). "Revolutions in RNA secondary structure prediction." J Mol Biol **359**(3): 526-32.

Mathews, D. H., M. E. Burkard, S. M. Freier, J. R. Wyatt and D. H. Turner (1999). "Predicting oligonucleotide affinity to nucleic acid targets." Rna **5**(11): 1458-69.

Mathews, D. H., M. D. Disney, J. L. Childs, S. J. Schroeder, M. Zuker and D. H. Turner (2004). "Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure." Proc Natl Acad Sci U S A **101**(19): 7287-92.

Mathews, D. H., J. Sabina, M. Zuker and D. H. Turner (1999). "Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure." J Mol Biol **288**(5): 911-40.

Mathews, D. H. and D. H. Turner (2002). "Experimentally derived nearest-neighbor parameters for the stability of RNA three- and four-way multibranch loops." Biochemistry **41**(3): 869-80.

Mathews, D. H. and M. Zuker (2004). Predictive methods using RNA sequences. Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins. Hoboken, N.J., John Wiley & Sons**:** 143-170.

McCaskill, J. S. (1990). "The equilibrium partition function and base pair binding probabilities for RNA secondary structure." Biopolymers **29**(6-7): 1105-19.

McGall, G. H. and J. A. Fidanza (2001). "Photolithographic synthesis of high-density oligonucleotide arrays." Methods Mol Biol **170**: 71-101.

Mergny, J. L. and L. Lacroix (2003). "Analysis of thermal melting curves." Oligonucleotides **13**(6): 515-37.

Mergny, J. L. and J. C. Maurizot (2001). "Fluorescence resonance energy transfer as a probe for G-quartet formation by a telomeric repeat." Chembiochem **2**(2): 124-32.

Merino, E. J. and K. M. Weeks (2003). "Fluorogenic resolution of ligand binding by a nucleic acid aptamer." J Am Chem Soc **125**(41): 12370-1.

Merino, E. J. and K. M. Weeks (2005). "Facile conversion of aptamers into sensors using a 2'-ribose-linked fluorophore." J Am Chem Soc **127**(37): 12766-7.

Michel, F., A. D. Ellington, S. Couture and J. W. Szostak (1990). "Phylogenetic and genetic evidence for base-triples in the catalytic domain of group I introns." Nature **347**(6293): 578-80.

Micura, R. and C. Hobartner (2003). "On secondary structure rearrangements and equilibria of small RNAs." Chembiochem **4**(10): 984-90.

Mignon, P., S. Loverix, J. Steyaert and P. Geerlings (2005). "Influence of the pi-pi interaction on the hydrogen bonding capacity of stacked DNA/RNA bases." Nucleic Acids Res **33**(6): 1779-89.

Morse, D. P. (2007). "Direct selection of RNA beacon aptamers." Biochem Biophys Res Commun **359**(1): 94-101.

Nagel, J. H. and C. W. Pleij (2002). "Self-induced structural switches in RNA." Biochimie **84**(9): 913-23.

Nimjee, S. M., C. P. Rusconi and B. A. Sullenger (2005). "APTAMERS: An Emerging Class of Therapeutics." Annu Rev Med **56**: 555-583.

Nussinov, R., Piecznik G., Griggs, J.R., and Kleitman, D.J. (1978). "Algorithms for loop matching." SIAM Journal of Applied Mathmatics **35**(1): 68-82.

Nutiu, R. and Y. Li (2003). "Structure-switching signaling aptamers." J Am Chem Soc **125**(16): 4771-8.

Nutiu, R. and Y. Li (2004). "Structure-switching signaling aptamers: transducing molecular recognition into fluorescence signaling." Chemistry **10**(8): 1868-76.

Nutiu, R. and Y. Li (2005). "Aptamers with fluorescence-signaling properties." Methods **37**(1): 16-25.

Nutiu, R. and Y. Li (2005). "In vitro selection of structure-switching signaling aptamers." Angew Chem Int Ed Engl **44**(7): 1061-5.

Nuwaysir, E. F., W. Huang, T. J. Albert, J. Singh, K. Nuwaysir, A. Pitas, T. Richmond, T. Gorski, J. P. Berg, J. Ballin, M. McCormick, J. Norton, T. Pollock, T. Sumwalt, L. Butcher, D. Porter, M. Molla, C. Hall, F. Blattner, M. R. Sussman, R. L. Wallace, F. Cerrina and R. D. Green (2002). "Gene expression analysis using oligonucleotide arrays produced by maskless photolithography." Genome Res **12**(11): 1749-55.

Osborne, S. E., I. Matsumura and A. D. Ellington (1997). "Aptamers as therapeutic and diagnostic reagents: problems and prospects." Curr Opin Chem Biol **1**(1): 5-9.

Padmanabhan, K., K. P. Padmanabhan, J. D. Ferrara, J. E. Sadler and A. Tulinsky (1993). "The structure of alpha-thrombin inhibited by a 15-mer single-stranded DNA aptamer." J Biol Chem **268**(24): 17651-4.

Pai, S., A. D. Ellington and M. Levy (2005). "Proximity ligation assays with peptide conjugate 'burrs' for the sensitive detection of spores." Nucleic Acids Res **33**(18): e162.

Pan, Q., X. L. Zhang, H. Y. Wu, P. W. He, F. Wang, M. S. Zhang, J. M. Hu, B. Xia and J. Wu (2005). "Aptamers that preferentially bind type IVB pili and inhibit human monocytic-cell invasion by Salmonella enterica serovar typhi." Antimicrob Agents Chemother **49**(10): 4052-60.

Patel, D. J. (1997). "Structural analysis of nucleic acid aptamers." Curr Opin Chem Biol **1**(1): 32-46.

Patel, D. J. and A. K. Suri (2000). "Structure, recognition and discrimination in RNA aptamer complexes with cofactors, amino acids, drugs and aminoglycoside antibiotics." Reviews in Molecular Biotechnology **74**(1): 39-60.

Penchovsky, R. and R. R. Breaker (2005). "Computational design and experimental validation of oligonucleotide-sensing allosteric ribozymes." Nat Biotechnol **23**(11): 1424-33.

Proske, D., M. Blank, R. Buhmann and A. Resch (2005). "Aptamers--basic research, drug development, and clinical applications." Appl Microbiol Biotechnol **69**(4): 367-74.

Rajendran, M. and A. D. Ellington (2007). "Selection of fluorescent aptamer beacons that light up in the presence of zinc." Anal Bioanal Chem.

Rankin, C. J., E. N. Fuller, K. H. Hamor, S. A. Gabarra and T. P. Shields (2006). "A simple fluorescent biosensor for theophylline based on its RNA aptamer." Nucleosides Nucleotides Nucleic Acids **25**(12): 1407-24.

Rivas, E. and S. R. Eddy (1999). "A dynamic programming algorithm for RNA structure prediction including pseudoknots." J Mol Biol **285**(5): 2053-68.

Rivas, E., R. J. Klein, T. A. Jones and S. R. Eddy (2001). "Computational identification of noncoding RNAs in E. coli by comparative genomics." Curr Biol **11**(17): 1369-73.

Robertson, M. P. and A. D. Ellington (2000). "Design and optimization of effector-activated ribozyme ligases." Nucleic Acids Res **28**(8): 1751-9.

Robertson, M. P., S. M. Knudsen and A. D. Ellington (2004). "In vitro selection of ribozymes dependent on peptides for activity." Rna **10**(1): 114-27.

Roth, A. and R. R. Breaker (2004). "Selection in vitro of allosteric ribozymes." Methods Mol Biol **252**: 145-64.

Rumbles, G., A. J. Brown and D. Phillips (1991). "Time-resolved evanescent wave induced fluorescence spectroscopy." J. Chem. Soc. Farady Trans. **87**(6): 825-830.

Rupert, P. B. and A. R. Ferre-D'Amare (2001). "Crystal structure of a hairpin ribozyme-inhibitor complex with implications for catalysis." Nature **410**(6830): 780-6.

SantaLucia, J., Jr. (1998). "A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics." Proc Natl Acad Sci U S A **95**(4): 1460-5.

Schultze, P., R. F. Macaya and J. Feigon (1994). "Three-dimensional solution structure of the thrombin-binding DNA aptamer d(GGTTGGTGTGGTTGG)." J Mol Biol **235**(5): 1532-47.

Seelig, G., D. Soloveichik, D. Y. Zhang and E. Winfree (2006). "Enzyme-free nucleic acid logic circuits." Science **314**(5805): 1585-8.

Shapiro, B. A., Y. G. Yingling, W. Kasprzak and E. Bindewald (2007). "Bridging the gap in RNA structure prediction." Curr Opin Struct Biol **17**(2): 157-65.

Silverman, S. K. (2003). "Rube Goldberg goes (ribo)nuclear? Molecular switches and sensors made from RNA." Rna **9**(4): 377-83.

Soukup, G. A. and R. R. Breaker (1999). "Engineering precision RNA molecular switches." Proc Natl Acad Sci U S A **96**(7): 3584-9.

Soukup, G. A. and R. R. Breaker (1999). "Relationship between internucleotide linkage geometry and the stability of RNA." Rna **5**(10): 1308-25.

Soukup, G. A., G. A. Emilsson and R. R. Breaker (2000). "Altering molecular recognition of RNA aptamers by allosteric selection." J Mol Biol **298**(4): 623-32.

Stojanovic, M. N., P. de Prada and D. W. Landry (2000). "Fluorescent Sensors Based on Aptamer Self-Assembly." Journal of the American Chemical Society **122**(46): 11547-11548.

Stojanovic, M. N., P. de Prada and D. W. Landry (2001). "Aptamer-based folding fluorescent sensor for cocaine." J Am Chem Soc **123**(21): 4928-31.

Stojanovic, M. N. and D. M. Kolpashchikov (2004). "Modular aptameric sensors." J Am Chem Soc **126**(30): 9266-70.

Stojanovic, M. N. and D. W. Landry (2002). "Aptamer-based colorimetric probe for cocaine." J Am Chem Soc **124**(33): 9678-9.

Suess, B., B. Fink, C. Berens, R. Stentz and W. Hillen (2004). "A theophylline responsive riboswitch based on helix slipping controls gene expression in vivo." Nucleic Acids Res **32**(4): 1610-4.

Tang, J. and R. R. Breaker (1997). "Rational design of allosteric ribozymes." Chem Biol **4**(6): 453-9.

Tang, J. and R. R. Breaker (1998). "Mechanism for allosteric inhibition of an ATP-sensitive ribozyme." Nucleic Acids Res **26**(18): 4214-21.

Tang, J., J. Xie, N. Shao and Y. Yan (2006). "The DNA aptamers that specifically recognize ricin toxin are selected by two in vitro selection methods." Electrophoresis **27**(7): 1303-11.

Tang, J., T. Yu, L. Guo, J. Xie, N. Shao and Z. He (2007). "In vitro selection of DNA aptamer against abrin toxin and aptamer-based abrin direct detection." Biosens Bioelectron **22**(11): 2456-63.

Tasset, D. M., M. F. Kubik and W. Steiner (1997). "Oligonucleotide inhibitors of human thrombin that bind distinct epitopes." J Mol Biol **272**(5): 688-98.

Tinoco, I., Jr. and C. Bustamante (1999). "How RNA folds." J Mol Biol **293**(2): 271-81.

Tinoco, I., Jr., O. C. Uhlenbeck and M. D. Levine (1971). "Estimation of secondary structure in ribonucleic acids." Nature **230**(5293): 362-7.

Tsiang, M., A. K. Jain, K. E. Dunn, M. E. Rojas, L. L. Leung and C. S. Gibbs (1995). "Functional mapping of the surface residues of human thrombin." J. Biol. Chem. **270**(28): 16854-63.

Tuerk, C. and L. Gold (1990). "Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase." Science **249**(4968): 505-10.

Tung, C. S., T. I. Oprea, G. Hummer and A. E. Garcia (1996). "Three-dimensional model of a selective theophylline-binding RNA molecule." J Mol Recognit **9**(4): 275-86.

Tyagi, S. and F. R. Kramer (1996). "Molecular beacons: probes that fluoresce upon hybridization." Nat Biotechnol **14**(3): 303-8.

Verma, S., S. Jager, O. Thum and M. Famulok (2003). "Functional tuning of nucleic acids by chemical modifications: tailored oligonucleotides as drugs, devices, and diagnostics." Chem Rec **3**(1): 51-60.

Voss, B., C. Meyer and R. Giegerich (2004). "Evaluating the predictability of conformational switching in RNA." Bioinformatics **20**(10): 1573-82.

Walter, A. E., D. H. Turner, J. Kim, M. H. Lyttle, P. Muller, D. H. Mathews and M. Zuker (1994). "Coaxial stacking of helixes enhances binding of oligoribonucleotides and improves predictions of RNA folding." Proc Natl Acad Sci U S A **91**(20): 9218-22.

Wang, J., Y. Jiang, C. Zhou and X. Fang (2005). "Aptamer-based ATP assay using a luminescent light switching complex." Anal Chem **77**(11): 3542-6.

Wang, K. Y., S. McCurdy, R. G. Shea, S. Swaminathan and P. H. Bolton (1993). "A DNA aptamer which binds to and inhibits thrombin exhibits a new structural motif for DNA." Biochemistry **32**(8): 1899-904.

Wang, S. and K. A. White (2007). "Riboswitching on RNA virus replication." Proc Natl Acad Sci U S A **104**(25): 10406-11.

Waterman, M. S. and T. F. Smith (1978). "RNA secondary structure: A complete mathematical analysis." Math Biosci **42**: 257-266.

Wu, L. and J. F. Curran (1999). "An allosteric synthetic DNA." Nucleic Acids Res **27**(6): 1512-6.

Wuchty, S., W. Fontana, I. L. Hofacker and P. Schuster (1999). "Complete suboptimal folding of RNA and the stability of secondary structures." Biopolymers **49**(2): 145-65.

Xayaphoummine, A., T. Bucher and H. Isambert (2005). "Kinefold web server for RNA/DNA folding path and structure prediction including pseudoknots and knots." Nucleic Acids Res **33**(Web Server issue): W605-10.

Xia, T., J. SantaLucia, Jr., M. E. Burkard, R. Kierzek, S. J. Schroeder, X. Jiao, C. Cox and D. H. Turner (1998). "Thermodynamic parameters for an expanded nearest-neighbor model for formation of RNA duplexes with Watson-Crick base pairs." Biochemistry **37**(42): 14719-35.

Yamamoto, R., T. Baba and P. K. Kumar (2000). "Molecular beacon aptamer fluoresces in the presence of Tat protein of HIV-1." Genes Cells **5**(5): 389-96.

Yamana, K., Y. Ohtani, H. Nakano and I. Saito (2003). "Bis-pyrene labeled DNA aptamer as an intelligent fluorescent biosensor." Bioorg Med Chem Lett **13**(20): 3429-31.

Yang, L., C. W. Fung, E. J. Cho and A. D. Ellington (2007). "Real-time rolling circle amplification for protein detection." Anal Chem **79**(9): 3320-9.

Ye, X., A. Gorin, R. Frederick, W. Hu, A. Majumdar, W. Xu, G. McLendon, A. Ellington and D. J. Patel (1999). "RNA architecture dictates the conformations of a bound peptide." Chem Biol **6**(9): 657-69.

Zimmermann, G. R., R. D. Jenison, C. L. Wick, J. P. Simorre and A. Pardi (1997). "Interlocking structural motifs mediate molecular discrimination by a theophylline-binding RNA." Nat. Struct. Biol. **4**(8): 644-649.

Zimmermann, G. R., C. L. Wick, T. P. Shields, R. D. Jenison and A. Pardi (2000). "Molecular interactions and metal binding in the theophylline-binding core of an RNA aptamer." Rna **6**(5): 659-67.

Zuker, M. (1989). "On finding all suboptimal foldings of an RNA molecule." Science **244**(4900): 48-52.

Zuker, M. (2000). "Calculating nucleic acid secondary structure." Curr Opin Struct Biol **10**(3): 303-10.

Zuker, M. (2003). "Mfold web server for nucleic acid folding and hybridization prediction." Nucleic Acids Res **31**(13): 3406-15.

Zuker, M. and D. Sankoff (1984). "RNA secondary structures and their prediction." Bull Math Biol **46**: 591-621.

Zuker, M. and P. Stiegler (1981). "Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information." Nucleic Acids Res **9**(1): 133-48.

Vita

Bradley Hall was born February 8th 1977 in Indianapolis, Indiana to Florence Gaylord and Arthur Hall III. His father moved to Cicero, Indiana where Brad graduated 7th in his class from Hamilton Heights High School in Arcadia, Indiana. Upon graduation, he entered Indiana University in Bloomington Indiana and sought a degree in Microbiology, attaining his B.S. in 1999. During college he worked in the laboratory of Jeffery Palmer researching phylogenetic relationships within the plant genus *Pelargonium*. In 1999, he moved to Austin, Texas to work with Andrew D. Ellington to develop nucleic acid taggants. Bradley entered graduate school at the University of Texas at Austin in 2002 where he focused on the computational and rational design of nucleic acid biosensors. During this time, he was awarded an IGERT fellowship. In addition to graduate work, he performed duties as the Ellington Lab system administrator and was provided the opportunity to teach as the interim Research Educator for Dr. Ellington's aptamer stream in the newly formed Freshman Research Initiative, which he plans to continue after graduation as a non-tenure track faculty member of The Unviersity of Texas at Austin. His publications include:

Hall, B., Hesselberth, J., Ellington, A.D. 2007. "Computational selection of nucleic acid biosensors via a slip structure model." Biosensors and Bioelectronics, 22 (9-10), 1939-1947.

Sokolov, K., Nida, D., Descour, M., Lacy, A., Levy, M., Hall, B., Dharmawardhane, S., Ellington, A., Korgel, B., Richards-Kortum, R. 2007. "Molecular optical imaging of therapeutic targets of cancer." Advances in Cancer Resarch, 96, 299-344.

Hall, B., Ellington, A.D. 2003. "Survival and polymerase chain reaction-based detection of nucleic acid taggant markers during bacterial growth and sterilization." Analytica Chimica Acta, 475, 67-73.

Yang, L., Fung, C., Hall, B., Cho, EJ., Levy, M., Ellington, A.D. 2007. "Real-Time RCA for Protein Detection." USP Ser. # 60/888,247. Patent Filed.

Permanent address: 870 Lanyard Dr. Cicero IN 46034

This dissertation was typed by the author.